Silent Networks: The Role of Inaccurate Beliefs in Reducing Useful Social Interactions*

Ronak Jain[†] Vatsal Khandelwal [‡]

December 1, 2025

Abstract

Can pessimistic beliefs about social norms create "silent" networks, where individuals rarely seek support despite benefits? A field experiment with informal workers shows that belief correction increases demand for social support by reducing perceived reputational costs. We use a network model to explain how misperceptions can arise and trap communities in a low-interaction equilibrium. Structural estimation predicts that the induced shift is too small to escape this equilibrium. Accordingly, two years later, those exposed to the treatment hold more optimistic beliefs, but below the true level conveyed. Counterfactuals show that stronger policies are required for persistence, though belief correction can increase the demand and funding to support them.

KEYWORDS: Social Networks, Social Norms, Beliefs, Risk Sharing, Mental Health.
JEL CLASSIFICATION: C93, D83, D85, D91, I12, I31, Z13.

*We are very grateful to Stefan Dercon, Marcel Fafchamps, Simon Quinn, and Alexander Teytelboym for their guidance and suggestions. We also thank Johannes Abeler, Emily Breza, Stefano Caria, Banu Demir, James Duffy, Benjamin Enke, Asim Khwaja, Gabriel Kriendler, Pramila Krishnan, Edward Glaeser, Matthew Jackson, Mahreen Mahmud, Kate Orkin, and Gautam Rao for their helpful comments. We thank seminar participants at various conferences including PacDev, NEUDC, the Econometric Society European Winter Meeting, the Workshop on Networks and Development, CSAE Conference, CEPR-TCD TIME Conference; OxDev (Oxford), and department workshops at Harvard University and the University of Oxford for their comments. Many thanks to Diana Beltekian, Binta Zahra Diop, Luke Milsom, Bassel Tarbush, Verena Wiedemann, and Hannah Zillessen for their feedback on previous versions of the paper. We are also deeply grateful to the NGO team, especially Bharati Chaturvedi, Ishani Palandukar, Shruti Sinha, and Balmukund Kumar for their collaboration and to Vikram Patel and Sadaf Vibha for sharing helpful information and resources regarding mental health and reviewing the appropriateness of the intervention in the local context. We extend sincere thanks to our field team, especially Preetham Rodrigues and Abdul Latheef at the Inchara Foundation, and Garima Tyagi, Swati Chauhan, Darshna Jadhav, Pushyami Chilakapati, Anshuman Singh, Dipti Pareek, and Aarti Malik, for help with surveying. Finally, we gratefully thank the Center for Study of African Economies, Oxford for funding the study. The project was approved by both the Harvard Institutional Review Board and CUREC, University of Oxford. The RCT was pre-registered at the AEA RCT registry: https://doi.org/10.1257/rct.6980-2.1.

†University of Zurich. Email: ronak.jain@econ.uzh.ch
‡University of Exeter. Email: v.khandelwal@exeter.ac.uk

1 Introduction

Individuals rely on their social networks to receive and provide support by sharing information, exchanging resources, and managing shocks (Beaman et al., 2021; Breza et al., 2019; Banerjee et al., 2024; Angelucci et al., 2018; Munshi and Rosenzweig, 2016). However, the effectiveness of social networks to perform these functions can be severely constrained if individuals do not *ask* for help. For instance, individuals may avoid discussing financial or mental-health concerns if they believe doing so could harm their reputation. Such beliefs can reduce the demand for support and affect the quality of risk-sharing within the community. Importantly, these beliefs about others may be inaccurate (Bursztyn and Yang, 2022).

This paper shows that inaccurate beliefs about peers can generate "silent networks": networks with limited interactions despite potential benefits. We combine evidence from a field experiment, survey experiments, and a structural model to show how communities can become stuck in a low-interaction, inefficient equilibrium, and that this can occur even when peers are in fact willing to engage. We study this among communities of informal workers in India that are characterized by high income volatility, limited access to formal support, and little policy attention (Marx et al., 2013). As 860 million people globally live in similar urban settlements (UN-Habitat, 2013), understanding why networks may fail to provide support even where they are most needed is an important question.

We conducted multiple rounds of data collection: a baseline survey in 2020, a randomized controlled trial in 2021, and additional larger-scale experiments in 2023. These data reveal four empirical patterns that motivate our study. First, individuals experience substantial income and consumption volatility. Second, they lack access to credit or professional support to manage financial or emotional distress. Third, despite limited formal support, most do not discuss financial or mental-health concerns with peers: 68% underestimate others' willingness to engage, with the majority doing so by more than 10 percentage points.[1] Fourth, these beliefs are consequential: individuals with more pessimistic beliefs interact less and have fewer peers for support within their networks. At the same time, having fewer peers in the borrowing network is associated with a higher probability of a consumption crisis and lower reported happiness.

Motivated by these patterns, we designed a field experiment that provides accurate information on peers' willingness to engage around financial and mental-health issues.[2] The intervention substantially increases engagement with community: treated individuals are 15 percentage points more likely to sign up for a savings group, 16 percentage points more willing to train as listening volunteers[3], and contribute 29% more to finance these services. They also show a higher willingness to seek formal mental-health support and discuss physical-health concerns. These sizeable treatment effects suggest that even moderate upward revisions in perceived norms can meaningfully increase network engagement.

[1]Survey evidence and additional experiments show that social desirability bias is not driving the high willingness to engage that we measure around these topics.

[2]A prediction exercise with a separate sample from the same setting confirms this is perceived as a strong information shock, predicted to increase engagement with the network.

[3]This is similar to the services that charities like Samaritans offer in the UK and US.

We conduct follow-up surveys two weeks and two years later. Two weeks later, treated individuals report substantially higher interaction around both financial and mental-health concerns: they are 19 percentage points more likely to have reached out to any peer about financial issues and contact, on average, one additional peer relative to the control group. Two years later, those exposed to the treatment hold significantly more optimistic beliefs about others' willingness to engage, report higher interaction intensity, and experience lower consumption volatility. The causal link between beliefs about others and interactions also generalizes to a larger post-pandemic sample of 800 individuals.[4]

To understand mechanisms, we use additional survey data and experiments. We first identify the types of social costs that discourage individuals from engaging with others. Using a hypothetical choice experiment, we test whether these costs stem from reputation concerns (gossip), signaling concerns (loss of job-related information), or interaction concerns (fear of insensitive responses). Individuals predict whether an advice-taking link on financial or mental-health matters would form between two randomly chosen community members, while we vary the characteristics of the hypothetical advisor. Across roughly 4,740 predicted links, reputation and interaction costs emerge as the main barriers, whereas signaling plays a limited role. Crucially, these costs are most salient for individuals with pessimistic beliefs, explaining why belief correction disproportionately increases their engagement.

We also address methodological concerns by implementing a list experiment and an experiment varying the visibility of participant responses to the surveyor, ruling out social desirability and experimenter demand. We further discuss that alternative explanations, such as updating beliefs about the benefits of interaction, the prevalence of distress, or social pressure, are unlikely to account for our findings.

Next, we adapt and structurally estimate a network model based on Jackson and Yariv (2007) to explain why individuals may systematically underestimate others' willingness to engage[5] and whether belief correction can shift equilibrium behavior. The key theoretical insight is that individuals can systematically underestimate others' willingness to engage because they infer the norm from the behavior of their peers, who—due to the "friendship paradox" (Feld, 1991; Jackson, 2019)—are more connected and therefore less likely to engage.[6] Intuitively, highly connected individuals face greater reputational risk when challenging norms (Young, 2015). This can trap communities in a Bayes–Nash equilibrium with persistently low dialogue. Belief correction can therefore help shift behavior in the short run, but whether these changes persist depends on the dynamics of best responses.

We structurally estimate the model using simulated method of moments combined with an equilibrium-selection criterion, leveraging the exogenous variation in beliefs generated by the RCT. The equilibrium that best fits the data is one of low engagement. The model predicts

[4] This addresses concerns that the findings are specific to that time period and the smaller sample sizes in the first two waves (350 at baseline and 180 in the intervention) due to the onset of the pandemic. In this validation exercise, participants receive information about *other* communities, so the underlying mechanism differs from the main RCT; see Section 5.7 and Appendix F for details.

[5] We use "engagement" following Jackson (2019), who study misperceptions in settings with positive peer effects.

[6] The friendship paradox states that an individual's friends, on average, have more friends than they do.

that the belief-shifting intervention generates only short-run changes in behavior and would be insufficient to shift the community to a new equilibrium. Consistent with this, two years later, those exposed to the treatment information hold significantly more optimistic beliefs, but their beliefs remain below the true levels conveyed during the intervention.

We then evaluate how difficult it would be for alternative interventions to convert the short-run effects of belief correction into a persistent equilibrium shift. Counterfactuals show that doing so would require large changes: perceived benefits of interaction would need to increase by roughly 50% of the estimated mean (by conducting awareness sessions, for example), or individual sensitivity to social norms would need to fall by about one-third (by setting up formal job information platforms, for example, so that individuals do not worry about reputation costs when asking for financial support). Such interventions are substantially more demanding than belief correction, especially in resource-constrained settings. Nonetheless, our results show that belief correction can raise both the demand and financial contributions for such intensive policies.

We make three contributions to existing literature. First, we contribute to the extensive literature on social networks by establishing a causal link between beliefs about social norms and the *demand* for network interactions.[7] While prior work has shown that trust, commitment, and enforcement constraints hinder cooperation (Ambrus and Elliott, 2021; Möbius and Rozenblat, 2016; Jackson et al., 2012; Karlan et al., 2009; Fafchamps and Gubert, 2007; Ligon et al., 2002), we show that inaccurate beliefs about norms can reduce network interactions even earlier in the process by preventing individuals from demanding support in the first place—a necessary precursor for any interaction. Moreover, while existing studies examine networks in rural settings[8], where caste, ethnic, or religious institutions structure social interactions, we study urban informal settlements where such institutions are weak, making individuals' beliefs central for enabling networks to function as social safety nets.

Second, we contribute to the growing literature on inaccurate beliefs and information provision experiments designed to correct misperceptions (e.g., Delavande (2023); Bursztyn and Yang (2022); Haaland et al. (2020); Bursztyn et al. (2020); Jackson (2019); Perkins et al. (2005, 1999)). We extend this work in two ways. First, we estimate a micro-founded explanation for why misperceptions arise, grounding them in the friendship paradox. Second, we document misperceptions in a domain with direct implications for social protection policy: individuals' willingness to interact with others on sensitive topics. The structural model allows us to predict the equilibrium effects of belief correction and to compare its effectiveness with alternative policies that do not target beliefs. In doing so, we also contribute to related work in social psychology, which documents that individuals underestimate others' willingness to interact, especially in low-stakes, one-off encounters (Epley and Schroeder, 2014; Epley, 2015). We extend this insight by showing that misperceptions can sustain a low-interaction community-level equilibrium, and that correcting them can increase social interactions in high-stakes real-world networks.

[7]See Breza et al. (2019); Breza (2016) for a review of this literature.

[8]See for example, Morten (2019); Banerjee et al. (2024); Munshi and Rosenzweig (2016); Banerjee et al. (2013); Munshi and Rosenzweig (2009)

Third, recent work shows that concerns such as signaling or shame can reduce incentives to seek advice or share information (Chandrasekhar et al., 2022, 2019; Banerjee et al., 2024). We show that belief correction can reduce the perceived reputational costs associated with violating social norms, offering a potential way to mitigate these constraints. Our results highlight that beliefs about others are malleable and may be easier to shift than deeper psycho-social factors typically targeted by more costly interventions. Moreover, belief correction can increase both the demand for and the financing of such costlier programs, providing a scalable pathway to relax these social barriers.

The rest of the paper proceeds as follows. Section 2 describes the context and presents the stylized facts motivating our experiment. Section 3 discusses belief elicitation and validation, and Sections 4 and 5 outline the experimental design and main results. Section 6 presents additional experiments and survey evidence on mechanisms. Section 7 introduces the model explaining why inaccurate beliefs may arise and reduce engagement, and Section 8 presents the structural estimation and policy counterfactuals. Section 9 concludes.

2 Context

Our sample consists of informal sector workers and their family members living in slums in Delhi, India. Most individuals work in waste picking, sorting, and recycling. We partnered with an NGO, Chintan, that promotes health and safety and advocates for the rights of waste workers. Globally, nearly 860 million people live in slums (UN-Habitat, 2013) and about 15 million earn a living through recycling waste in developing countries (Medina, 2008). This context is particularly relevant given the limited policy recognition and formal assistance available to these communities, making social networks critical as safety nets (Marx et al., 2013; Chaturvedi et al., 2018; Ivaschenko et al., 2018).

Tables 2 and 3 present descriptive statistics from the baseline survey with 352 individuals across 14 locations in Delhi (Figure A.1) in 2020 . Participants are, on average, 34 years old; 35% are female; and 67% are currently employed. On average, individuals take advice from about three peers in their "advice network", defined as the number of ties with whom they discuss mental-health concerns. They interact with roughly four peers overall, including for advice-taking, borrowing, lending, and work-related interactions (their 'overall network'). Individuals have lived in their current locations for an average of 20 years. Only 27% are migrants, and among them, just 18% frequently communicate with contacts in their origin locations. Finally, 72% report relying on their local networks for job information.

We conducted an additional survey with 791 individuals in the same setting in 2023 to ensure that the patterns are not specific to the pandemic period. Using data from both waves, we highlight three key features that motivate our intervention.

A. Financial & Psychological Distress and Lack of Formal Assistance: Individuals in our sample have very low and volatile incomes. Nearly 45% earn \$2.5–5 per day, and 35% earn less than \$2 per day. Income fluctuations are substantial: as Figure A.2 shows, the difference between the highest and lowest income over the previous six months averages 50% of the mean income. Nearly three in four individuals report facing serious financial difficulties—such as inability to afford a healthy diet, health expenditures, or children's education—often or very often in the past six months. Formal assistance is scarce: about half lack a bank account, and among those with one, 80% struggle to access credit. We also document high levels of psychological distress. Roughly 50% report frequently feeling overwhelmed by difficulties, consistent with the elevated stress levels reported in Table 2. Formal mental-health support is rarely accessed: roughly 90% of individuals in the control group report not feeling comfortable visiting a professional.

B. Low Levels of Interactions with Social Ties: Despite low access to formal support, dialogue around mental health and financial issues with peers is limited. Figure A.3 shows that nearly half the sample rarely discusses either topic with their social contacts. Network measures reveal a similar pattern. In Figure 1, the overall network degree distribution stochastically dominates the advice network degree distribution: individuals interact with more peers in general than they do for support on sensitive topics. While not all contacts are expected to provide support, the gap is quite large, consistent with network interactions being stuck at suboptimal levels. As we discuss in Section 3, these gaps are especially higher among individuals with severely pessimistic beliefs about others' willingness to engage. Similar patterns emerge for borrowing and lending links in the additional sample.

C. Networks can be beneficial: The limited use of network interactions is particularly notable given evidence that they can improve well-being. Individuals with more borrowing–lending ties report facing fewer consumption crises.[9] Each additional borrowing–lending connection is associated with a significant three–percentage-point reduction in the likelihood of experiencing a consumption crisis. An extra borrowing–lending tie is associated with a five–percentage-point lower probability of reporting below-median happiness, while an additional mental-health tie is associated with a 3.5–percentage-point higher probability of being above the median.[10] Although these associations are not causal, they are indicative of the important role networks can play in this setting.

3 Belief Elicitation and Validation

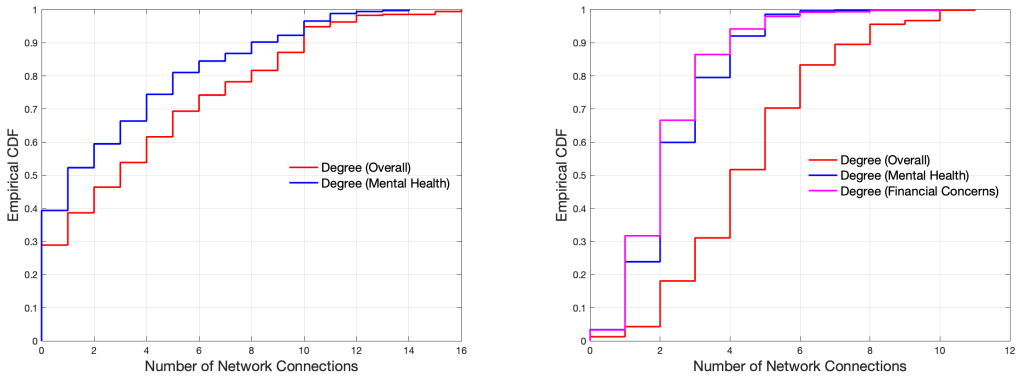
3.1 Belief Measurement

Measuring willingness to engage: To measure beliefs in a standardized way, we first clarified what we mean by physical and mental health. We then asked individuals about their willingness to engage with others on physical health, mental health, and finance/employment-related issues. 'Willingness to engage' was explained as the willingness to discuss ways to overcome concerns around these topics and how they may be preventing individuals from achieving their goals. This ensures we capture meaningful dialogue rather than casual conversation.

[9] A consumption crisis equals 1 if the respondent reports lacking sufficient resources for basic dietary, education, or health expenditures very often in the past six months.

[10] Similar positive correlations hold for interaction frequency.

Figure 1: Degree Distributions of Overall, Advice-Taking, and Financial Networks



Notes: The figure on the left plots the CDF of the number of connections in the overall and mental health-related advice networks in the main sample (in 2020) and the figure on the right additionally plots the CDF for borrowing-lending networks in the additional sample (in 2023). Individuals were asked to list up to ten other individuals in their community whom they interact with in any capacity i.e., to borrow/lend, take/give advice, work with etc. ("overall networks"), whom they only take advice from regarding mental health concerns ("advice-taking network"), and whom they contact for financial concerns ("financial concerns network").

Our measure does not assume a direction of help—whether the individual is giving or seeking support—but instead captures the underlying decision to enter a conversation at all. In practice, giving and seeking support are interdependent: offering help matters only once others are willing to speak, and seeking help matters only if others are willing to listen. Distinguishing the two would require assuming which side must move first to break the low-dialogue equilibrium, and our design intentionally avoids this. We interpret this measure as capturing the first stage of interaction: the willingness to entertain a conversation on a topic, which governs whether individuals initiate or engage in a dialogue.

Incentivized elicitation of beliefs about others' willingness to engage: We then elicited respondents' beliefs about their community's willingness to engage on three topics: physical health, mental health, and financial issues. We clarified what we mean by *community*, defining it as individuals living around the NGO center and affiliated with the NGO.[11] For example, to measure beliefs about engagement around financial concerns, we asked how many out of any 10 community members would be willing to share and discuss financial or employment-related problems with their friends, explore ways to overcome them, and reflect on how these issues might be affecting their goals.[12] Beliefs were elicited in an incentive-compatible way. Respondents received Rs.50 (one-third of the participation incentive) if their guess was accurate within 1 on the 0–10 scale.

Why do we measure beliefs about the community? We intentionally measure beliefs about the *community's* willingness to engage rather than beliefs about one's immediate friends. This captures perceived social norms and what individuals view as acceptable to discuss—even with close friends. People may interact within small peer groups, yet still fear gossip, information spreading, or reputational loss if such conversations violate broader community norms. Beliefs about community openness are therefore informative about these costs and

[11] Participants were familiar with this definition because the NGO is highly active in these areas.
[12] We used this format rather than percentages to ensure comprehension, based on pilot evidence.

help explain avoidance of topics seen as taboo. As we show later, these beliefs strongly correlate with individuals' own engagement decisions. Moreover, in a separate sample from the same setting, two-thirds of respondents predicted that engagement would increase if others learned that most community members were open to dialogue, underscoring that individuals view community-level norms as relevant for individual behavior.

Incentivized elicitation of beliefs about stress and stigma: To shed light on mechanisms, we also elicited two additional beliefs. First, we asked individuals to predict the average level of stress in their community. Specifically, they reported how many individuals among any 10 community members would 'often' or 'very often' feel that "difficulties were piling up so high that they could not overcome them." Comparing this prediction with the actual reports allows us to assess whether individuals overestimate or underestimate the prevalence of mental health distress. Similarly, to measure stigma, we asked respondents if they agreed with the statement: "People should stay away from individuals who have mental health issues," and asked them how many community members would agree with the statement.

3.2 Evidence of Misperceptions

Table 3 provides summary statistics for the elicited beliefs. About 71% of individuals report being willing to engage in dialogue around financial well-being, yet they expect only about 60% of their community to do so. Similarly, 63% say they are willing to discuss mental health, but believe only about 50% of others in their community would. Figure 2 shows the distribution of these beliefs.

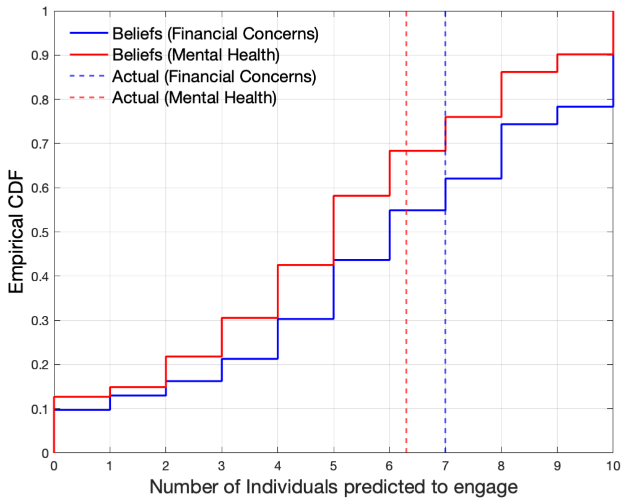
We find substantial misperceptions about others' willingness to engage. As shown in Figure A.5, 68% of participants underestimate the willingness to discuss mental-health in their own community, and 58% underestimate engagement around financial concerns. In contrast, most individuals overestimate the prevalence of stress in the community.

Magnitude of Misperceptions: Although average gaps appear modest, underestimation of willingness to discuss mental health exceeds 10 percentage points for most underestimators. Even small misperceptions can meaningfully raise the perceived social costs of engaging on sensitive topics because individuals tend to fear worst-case reactions—such as rejection or insensitivity—which depend on perceived norms. Thus, perceived costs can respond nonlinearly to beliefs, and modest misperceptions can produce large behavioral differences.

Social Desirability: We do not expect that the high willingness to engage reported is driven by social desirability. First, around 49% of the participants report agreeing that one must stay away from those with mental health issues, indicating that respondents felt comfortable expressing potentially stigmatizing views.[13] Second, uniformly express willingness to engage across topics: for example, 42% of those unwilling to engage on mental-health concerns are nevertheless willing to engage on financial matters. Third, as we show in the next section, willingness to engage is systematically correlated with baseline beliefs about peers' willingness to engage. Taken together, these patterns suggest that respondents answered

[13]Discussions with NGO staff and a psychologist suggest that the phrasing may have led participants to think of severe mental-health disorders.

Figure 2: Beliefs about Community Engagement



Notes: The solid lines in the figure plot the CDF of how many community members out of a randomly chosen 10 individuals believe will be willing to engage around financial concerns and mental health issues respectively. The dotted lines indicate the actual sample willingness to engage.

thoughtfully and that these measures are not driven by cheap talk. Additional experiments discussed in Section 6 further resolve this concern.

3.3 Validation of Elicited Beliefs

To assess the accuracy of elicited beliefs, we correlate them with individuals' past behavior, network characteristics, and economic outcomes, following Delavande et al. (2011). This allows us to confirm that the misperceptions we observe are not driven by measurement error. Across all checks, the correlations have the expected signs and support the validity of our measures.

Correlations of Beliefs and Individual Willingness to Engage: Participants' beliefs about their community's willingness to engage on mental, financial, and physical health are positively correlated with each other (Table A.1). This implies that optimism about the openness of the community on one dimension implies optimism around other dimensions as well. These beliefs are also correlated with an individual's own willingness to engage, suggesting that they capture meaningful variation in perceived norms and behavior.

Correlations of Beliefs and Frequency of Dialogue Intensity: Beliefs are strongly associated with the frequency of actual conversations. Regressing conversation frequency on beliefs about engagement shows that more optimistic beliefs predict higher dialogue around both mental health and financial issues (Table A.2). Panel B further shows that individuals who underestimate community engagement are significantly less willing to engage in dialogue around mental health themselves.

Correlations of Beliefs and Network Gaps: We also find that pessimistic beliefs are associated with structural features of individuals' social networks. Individuals with lower perceived community openness have larger gaps between their overall and advice networks

(Table A.2). These gaps—measured as the proportional difference between overall network degree and advice network degree—are especially large among "severe underestimators" (bottom 25th percentile of belief accuracy).

Correlations of Beliefs with Consumption Outcomes: Using data from the additional sample, we find that more pessimistic beliefs are correlated with higher consumption volatility and a greater likelihood of experiencing a consumption crisis (Table A.4). This highlights the behavioral and economic relevance of misperceptions.

Large-scale validation in 2023: To ensure that observed misperceptions are not specific to the COVID-19 period, we elicited beliefs in a larger sample with 791 individuals in 2023. Underestimation of community engagement remains high, and willingness to engage is similarly high across topics. As before, more optimistic beliefs are associated with having more mental-health and financial network ties (Table A.4). Full results are presented in Appendix Section F.

4 Experiment

4.1 Timeline and Design

The main experiment and endline surveys were conducted with 180 individuals from February to April 2021, with roughly half in the treatment group and half in the control group. In 2023, we implemented a larger-scale experiment with approximately 800 additional participants.[14]

The treatment group received two pieces of information at two different points during the survey: (i) the true average sample willingness to engage in dialogue about mental health, and (ii) information that the sample's willingness to engage in a dialogue about financial well-being was high.[15] More specifically, the statements used were:

Information 1: *"Just like we surveyed you, we also surveyed other people and we have found from their responses that X out of any 10 individuals in your community and communities similar to yours (affiliated with the NGO) are willing to discuss their mental health concerns with their peers, how they can be preventing them from achieving their goals, and how they can be overcome."*

Information 2: *"Just like we surveyed you, we also surveyed other people and we have found from their responses that the proportion of individuals willing to discuss their financial/work-related concerns with their peers, how they can be preventing them from achieving their goals, and how they can be overcome is high."*[16]

The second piece of information was provided closer to the end of the survey, before the

[14] Further details are provided in Section F of the appendix.

[15] We provided information about sample-level averages computed using data on willingness to engage from our initial surveys of about 350 individuals rather than community-level averages. We did this because some communities had low response rates, and we did not wish to shift beliefs about their community based on information obtained from small samples. Importantly, since the communities are very similar to each other, information from the entire sample is informative and helpful.

[16] Note that we do not provide a precise estimate here (the precise estimate of the proportion of individuals willing to share these concerns is 70%) because unlike mental health, where there is a higher proportion of under-estimators, in this case, 58% underestimate, and we do not wish to make them more pessimistic.

questions related to financial outcomes. All other components of the survey were identical for the control and treatment groups.

Out-of-sample predictions about the intervention's effects: We expect the treatment to function as a meaningful information shock, given limited dialogue around these topics and the stigma associated with discussing mental well-being. To assess whether individuals themselves viewed the information as relevant for behavior, we asked participants in our additional sample to predict how others would respond if they received Information 1. A large majority—67%—anticipated that engagement with savings groups would increase, and 42% expected this increase to be large.

4.2 Balance Tests

We show that the sample is balanced across control and treatment groups across a host of baseline variables including demographic information, network connections, dialogue intensity, beliefs about mental health, physical health, financial concerns, participant well-being, and own willingness to engage. In particular, we test balance on 44 variables and find that the sample is unbalanced only for 2 variables. These results are presented in Tables B.1 and B.2. We also regress treatment status on these baseline covariates and find that the F statistic is 0.68 and the corresponding p-value is 0.89. This suggests that the baseline characteristics are balanced across treatment and control. However, for robustness, we also run a specification that controls for any unbalanced covariates.

4.3 Specification

We regress each outcome on the treatment indicator for the individual, with robust standard errors. We additionally report p-values of wild bootstrapped t-statistics in line with Cameron et al. (2008) where we cluster the standard errors at the level of the NGO center. Further, we also conduct correction for tests of multiple hypotheses as per Benjamini et al. (2006); Anderson (2008). These q-values are computed at the level of the outcome families. We also run an additional robustness exercise where we take a very conservative approach and treat all the outcomes as one family before computing the q-values.

Finally, we also account for the fact that participants may not have always completed the endline survey as a result of which some outcomes may have received more responses than others. To address any balance-related concerns that may arise due to this, we adopt the following additional strategy as a robustness check in the appendix. We run balance tests for *each* sub-sample for which the outcome variable is non-missing. Then, we include the unbalanced controls (at 5% significance) in a conservative, robust specification. All the main results are robust to this alternative specification unless specified otherwise.

5 Results

We now present the effects of the intervention on pre-specified outcomes measured immediately after the treatment. These include (1) willingness to engage with the community on mental health and financial issues, (2) demand for formal support, (3) demand for additional information, (4) willingness to talk about other health-related issues, and (5) financial self-efficacy and stigma. Most outcomes were measured after Information 1; we indicate which outcomes were collected after both information statements.

5.1 Community Engagement

Community engagement is defined as the mean of three binary variables capturing willingness to have useful interactions with the community: (a) willingness to train as a "listening volunteer" for the community[17] (b) willingness to contribute financially to establish this listening service, and (c) willingness to participate in an NGO-facilitated savings group to pool savings to help each other in difficult times. The first two outcomes were collected after Information 1; the third was collected after both information statements.

As Table 4 shows, the intervention increases community engagement by 14 percentage points.[18] Figure 3 plots the treatment effects separately. Treated individuals are 16 percentage points more likely to enlist as volunteers (significant at 5%) and 12% more likely to financially contribute to the training. Their average contribution increases by Rs. 6.6 more, which is a 29% increase compared to the control group (significant at 5%).[19] We also find that while 67% of the control group is willing to participate in savings groups, the treatment increases this further by about 15 percentage points (significant at 5%).

These results indicate that correcting misperceptions about peers meaningfully increases individuals' willingness to interact with their networks—both through volunteering time and financially contributing to informal support avenues. All effects remain robust to the inclusion of unbalanced controls.

5.2 Demand for Formal Support

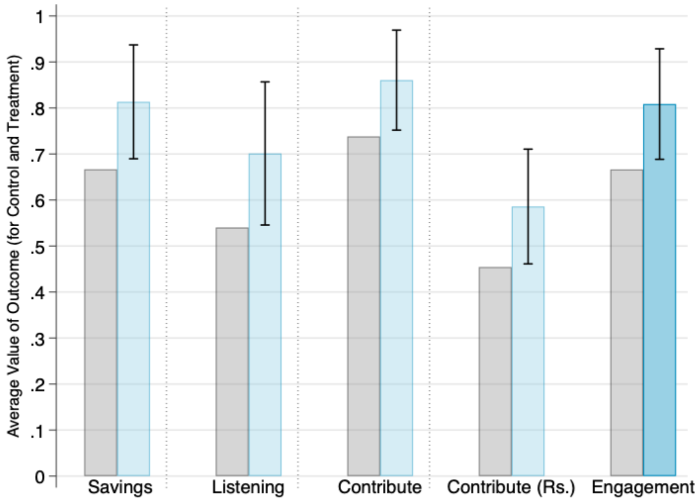
Panel B of Table 4 shows that the intervention also affects uptake of formal support. Treated participants report lower hesitation to speak with a doctor about mental health concerns, which is consistent with a reduction in perceived stigma or perceived costs of violating the social norm. At the same time, treatment reduces participation in experimenter-run depression scoring (PHQ-9) and reduces willingness to listen to government helpline information, both measured after both information statements. These patterns suggest that individuals may consider formal and informal support as substitutes: greater comfort with community-based dialogue may reduce interest in externally provided resources. All effects remain robust to adding unbalanced controls.

[17] This captures training to become a "mental health buddy" who listens to the worries of other community members, similar to the services provided by charities like Samaritans.

[18] All q-values are also significant at 5%.

[19] Participants were informed that their contributions would be deducted from a Rs 50 prize for those who won based on their accuracy of guesses. Baseline beliefs were balanced, so selective pessimism about winning is unlikely.

Figure 3: Treatment Effects on Engagement with the Network



Notes: The above figure shows the average value of each outcome for the control group (in gray) and the treatment group (in blue) with 95% confidence intervals for the difference between the two values. "Savings", "Listening", and "Contribute" refer to binary variables indicating whether the individual is willing to participate in savings groups, train and volunteer for a listening service, and actually contribute to set it up. "Engagement" is an average of these three variables. "Contribute (Rs.)" is the actual monetary contribution made by the individuals, normalized to be between 0 and 1.

5.3 Other Types of Dialogue

Panel C of Table 4 provides evidence for sizable positive spillover effects: treated individuals are 21% more likely to report willingness to discuss physical health-related issues with peers. Given baseline stigma around discussing physical health, this suggests that shifting beliefs about norms in one domain (mental health) can generalize to other sensitive domains. The result is robust to unbalanced controls.

5.4 Demand for Additional Information

Demand for additional information includes willingness to attend a one-hour NGO-led mental-health information session and willingness to listen to guidance on initiating sensitive conversations. The latter outcome was measured after participants received both information statements.

Table C.1 shows that the treatment does not affect participants' demand for additional information about mental health. Admittedly, our ability to detect effects on the take-up of the information session is limited due to the high baseline take-up of the information session as 67% of the control group agrees to participate. Anecdotally, we learnt that respondents perceived the session as similar to routine NGO trainings rather than as a stigmatized activity. We also find no effect on demand for guidance on having sensitive conversations.

5.5 Other Outcomes

We do not detect treatment effects on financial self-efficacy or stigma (Supplementary Appendix Table C.1). Financial self-efficacy measures confidence in (a) managing financial or employment-related concerns in the coming months, (b) financing children's education expenses, and (c) starting a small business. We do not find any effect on any of these outcomes. While one might expect belief correction about peers' willingness to engage on both financial and mental-health issues to affect these outcomes, the null results are consistent with the lack of existing financial risk-sharing: optimism about dialogue need not immediately translate into confidence about borrowing or financial support if such links are absent.

Stigma is measured using: (i) willingness to have one's name publicly shared as a participant in the information session, (ii) a list-experiment measure of agreement with stigmatizing statements, and (iii) willingness to disclose depression scores to the surveyor (conditional on participating in the depression scoring). We do not find any treatment effects on these outcomes. Notably, 92% of control participants were willing to have their names shared, suggesting this was a weak measure of stigma—likely because NGO-run information sessions are common and not perceived as stigmatizing.

5.6 Two-week and Two-year Follow-Up

5.6.1 Effects after two weeks

We conducted a follow-up survey two weeks after the intervention in 2021. As described in Appendix Section E.1, we pre-specified all follow-up outcomes. Due to the peak of the COVID-19 crisis in India during April 2021, we were able to reach only 112 respondents (57 treated). Despite the smaller sample, balance checks on 44 baseline variables show imbalance on only 2, and robustness checks controlling for these variables confirm that attrition is not correlated with treatment status or baseline characteristics.

Table 5 shows that treated individuals are 11–19% more likely to have reached out (via phone or SMS) to at least one peer to discuss their own or others' mental-health or financial concerns in the two weeks following treatment. Treated respondents also contacted roughly one additional peer to discuss both mental-health issues ($p < 0.05$) and employment-related concerns ($p < 0.01$). The effects are stronger for initiating contact than for receiving it, consistent with the fact that only treated individuals received the information. Additional results on well-being and dictator games, as well as a discussion of potential COVID-related concerns, are presented in Appendix Section E.

5.6.2 Effects after two years

In 2023, we conducted a demographic survey in the same setting and classified individuals into four categories: (1) *Previous Participant* — reported participating in the 2020–21 surveys/experiment; (2) *Heard Community Beliefs* — reported hearing information about their community's willingness to engage on mental-health and financial issues; and (3) *Heard Community Beliefs + Previous Participant*— both participated earlier and reported hearing the information; and (4) *None* of the above.[20] Although we cannot directly match individuals

[20] This survey was conducted as part of the additional surveys completed in 2023 where we had an overlap of 1 center from the previous survey which we use to measure long-term effects.

across waves, these self-reports enable us to identify likely exposure. Of 469 respondents, 60 fall into Category 1, 37 into Category 2, and 24 into Category 3. We acknowledge that exposure is a proxy and may capture differences in individual characteristics apart from exposure to treatment. We therefore use post-double-selection Lasso (Belloni et al., 2014) to adjust for observable differences such as gender, income, age, and overall network degree.

Tables 6–7 report the associations between these exposure categories and outcomes in 2023. Individuals in Categories 1–3 consistently hold more optimistic beliefs about others' willingness to engage on mental-health and financial concerns, with effects strongest among those reporting that they heard the information.[21] They are also more willing to engage themselves, more likely to report greater engagement in dialogue (i.e., above-median) with peers, and experience lower consumption volatility and fewer consumption crises. These patterns suggest that beliefs about peers can influence network interactions and, in turn, economic outcomes.

While the evidence is only suggestive, it is notable that many relationships remain significant even under the most conservative measure of exposure (being a previous participant). However, it appears unlikely that beliefs and engagement have shifted permanently. For instance, among individuals reporting that they heard the information, average beliefs remain below the true value conveyed during the intervention. We discuss whether the intervention could change equilibrium beliefs through a structural model in Section 8.

5.7 Larger-scale Validation

We briefly summarize the results from an additional experiment conducted in 2023 with roughly 800 individuals across two NGO centers. This exercise tests whether (a) the patterns of low dialogue and underestimation of peer willingness to engage, and (b) the causal effect of beliefs on individuals' own engagement, are specific to the pandemic. We find that the same baseline patterns of low dialogue and pessimistic beliefs also hold in this larger sample.

In addition, we implemented a variation of the intervention in which participants received information about the willingness to engage in *other* similar communities, rather than their own. Namely, participants were shown information collected previously from communities in 2021.[22] We find that this information continues to have a significant causal impact on willingness to engage. However, the mechanisms differ from the main experiment. Individuals who are pessimistic about their own community become *less* likely to join a savings group when told that *other* communities are more willing to engage, whereas those who are optimistic about their own community respond in the opposite direction. A similar pattern arises for volunteering in listening services. Treated participants continue to make higher financial contributions to support informal interaction avenues, particularly among those who are pessimistic about the social norm.

These results indicate that the treatment effect depends strongly on the reference group: information about one's own community shifts beliefs about peers directly, whereas

[21] We drop respondents who reported hearing community views as part of the current 2023 surveys.

[22] This was owing to financial and logistical constraints of conducting belief elicitation again.

information about other communities induces relative comparisons. A detailed discussion of these results appears in Appendix Section F.

5.8 Heterogeneity

We examine heterogeneity in treatment effects by baseline willingness to engage, dialogue intensity, network size, stigma, and beliefs.[23] The results are shown in Tables C.2-C.7 in the appendix. Below, we summarize the main patterns.

5.8.1 Heterogeneity by Baseline Engagement

Table C.2 shows that the treatment effects on community engagement are substantially larger among individuals who were initially unwilling to engage on mental-health concerns. For these respondents, the intervention increases willingness to (a) engage with their community, (b) attend a mental-health information session, and (c) listen to good practices for sensitive conversations.[24] Similarly, Table C.3 shows that effects are significantly larger for individuals with below-median baseline dialogue on mental-health and financial issues. This pattern is consistent with the intervention benefiting those who were least engaged at baseline, therefore helping those who need it most.

5.8.2 Heterogeneity by Baseline Networks

Table C.4 shows strong heterogeneity by the overall network size. More connected individuals exhibit larger treatment effects: each additional peer raises the probability of participating in a savings group by roughly 3 p.p., while treatment effects for individuals with zero peers are effectively zero. More connected individuals also contribute significantly more to establishing the listening service in response to treatment, reversing the baseline pattern in which they are less likely to contribute. We return to this result in Section 7, where the theoretical framework predicts larger gains for highly connected individuals.

5.8.3 Heterogeneity by Baseline Stigma

As shown in Table C.5, individuals expressing mental-health related stigma (those agreeing that they should "stay away" from people with mental-health issues) donate Rs. 7 less to the listening service than those without stigma. However, among individuals who express stigma, treated participants donate Rs. 9 more than their control counterparts. While information on community openness does not directly reduce stigma, it substantially increases engagement among those who initially hold stigmatizing views—potentially because pessimistic beliefs about norms and expressed stigma are correlated, and the intervention relaxes perceived social costs of engagement.

[23] Heterogeneity by baseline beliefs about stress is discussed with mechanisms in Section 6.

[24] For the last measure, the interaction term is marginal, but the bootstrapped p-value is significant at 5%.

5.8.4 Heterogeneity by Baseline Beliefs

Finally, Table C.6 examines heterogeneity by whether individuals underestimated community engagement around mental-health and financial issues. Our ability to detect effects is limited, but effect sizes are generally larger for underestimators, consistent with the mechanism that belief correction matters most for those holding pessimistic priors.

6 Mechanisms

6.1 Main Mechanism: Lower Perceived Costs of Violating the Social Norm

To understand the mechanisms underlying the treatment effects, we use a survey experiment with our additional sample of around 800 individuals. Each individual was asked to predict whether a link would form between two hypothetical community members, where agent A is a potential advisor and agent B needs support for financial or mental health-related concerns. We exogenously vary three characteristics of advisor A's: (1) Network centrality (testing reputational costs—concerns about gossip), (2) Contacts in private jobs (testing signaling costs—fear of negative inferences affecting job referrals), and (3) Sensitivity training attendance (testing interaction costs—fear of being mocked or treated insensitively). Each respondent evaluated three randomly generated profiles, yielding 4,740 total predictions across mental-health and financial scenarios.

Overall, 23% of mental-health links and 24% of financial links were predicted not to form—reflecting meaningful perceived barriers even in a hypothetical setting. Table 8 shows that individuals with more pessimistic beliefs about community engagement are significantly more likely to predict that links will not form, especially when A is highly central or has attended sensitivity training. This pattern is consistent across both OLS and Probit specifications.

To understand why, we asked respondents to explain their answers. Among those rejecting links when A is central, 42% cited concerns that A would not have time, and 39% cited fear of gossip. Among those rejecting links when A had job contacts, 30% feared being judged as an unsuitable candidate for referrals—consistent with signaling concerns, even though this characteristic did not causally reduce predicted link formation on average. Surprisingly, sensitivity training attendance acted as a negative signal: respondents interpreted participation as suggesting that "something must be wrong" with A, reinforcing interaction-related concerns.

Taken together, these results indicate that reputational and interaction costs—both shaped by perceived social norms—are key barriers to forming supportive ties. Importantly, these costs are substantially higher among those with more pessimistic beliefs about others' willingness to engage, precisely the group for whom the belief correction treatment had the largest effects. This interpretation is reinforced by qualitative interviews: participants reported fear of being shamed, mocked, or gossiped about if their request for help was rejected, describing themselves as "ashamed," "worried they will make fun later," and "suffocated or trapped because others can make fun of their poverty."

6.2 Alternative Mechanisms

We next use additional experiments and survey evidence to discuss alternative explanations for the treatment effects.

6.2.1 Social Desirability and Experimenter Demand

We assess whether our findings could reflect social desirability or experimenter-demand effects using (a) a list experiment and (b) random variation in whether key responses were visible to the enumerator.

List experiment: Individuals were randomly assigned to a control group or to one of two treatment groups that received either one or two statements about mental health-related and financial well-being-related engagement with peers, respectively.[25] Respondents only reported how many statements they agreed with, masking agreement with any specific item. If social desirability drove reported willingness to engage, we would expect no difference across groups. Instead, agreement rises systematically with the inclusion of engagement-related statements (2.21 in control; 2.97 and 3.71 in treatment groups), and the differences are statistically significant.

Enumerator-visibility experiment: In a separate exercise, we randomized whether willingness to join a savings group or train as a listener was reported privately or aloud to the enumerator. Responses are indistinguishable across the two conditions and do not interact with treatment status (Table D.2), indicating that neither social desirability nor experimenter demand explains the results.

Additional evidence from costly outcomes: The treatment raises costly behaviors and incentivized outcomes—such as financial contributions—which are unlikely to be driven by cheap talk. Moreover, the largest effects occur among individuals with low baseline engagement and low baseline dialogue (Tables C.2–C.3), which is inconsistent with a simple desirability story.

6.2.2 Updating Beliefs about the Incidence of Stress

The treatment could also affect behavior if individuals revise beliefs about how stressed or burdened their peers are. However, baseline data show that over 60% overestimate community stress levels, making it unlikely that lack of awareness of stress is a key constraint. Consistent with this, treatment effects do not vary by whether respondents over- or underestimate stress (Table C.7). This suggests that changes in perceived incidence of stress are not the main pathway.

[25] The statements are as follows: (1) Only individuals who have received formal education should enter into politics. (2) The Delhi government is taking the required steps to deal with air pollution. (3) Teachers should be paid more remuneration than film actors. (4) Individuals should take time to listen to the mental health concerns of their peers (Only to Treatment Groups 1 and 2) (5) Individuals should take time to listen to the employment or money-related concerns of their peers. (Only to Treatment Group 2).

6.2.3 Altruism or Social Pressure

Could individuals increase engagement out of altruism or a desire to appear supportive? Key patterns suggest this is unlikely. Altruism alone is hard to reconcile with the low observed dialogue at baseline, despite widespread awareness of high stress. Second, the treatment *reduces* take-up of depression scoring and helpline information—consistent with substitution between community support and formal support, and not with individuals engaging purely out of altruistic concern for others. Third, contributions to the listening service were explicitly private, and willingness to have one's name shared publicly does not differ by treatment status. Together, these patterns are suggestive that social pressure or altruism are unlikely to be primary mechanisms.

6.2.4 Updating Beliefs about Benefits of Interacting

Another possibility is that individuals update beliefs about the benefits of interacting rather than the costs. However, baseline data show that willingness to engage is already high, even among individuals who rarely engage, suggesting that perceived benefits are not the binding constraint. Moreover, while the frequency of dialogue is low at baseline, such interactions are not absent altogether; respondents could observe any existing engagement within their community—for instance, borrowing–lending ties—so it seems less plausible that they are unaware of the potential benefits. Qualitative interviews also confirm that respondents view such interactions as valuable. These patterns point to perceived social costs—not a lack of perceived benefits—as the more plausible barrier.

7 Theoretical Framework

We adapt the frameworks in Jackson and Yariv (2007) and Jackson (2019) to show how individuals in a fixed network form beliefs about others' willingness to engage and how these beliefs shape their own decisions to interact around sensitive topics. The model illustrates how misperceptions naturally arise and motivates our information intervention.

Let N denote the set of individuals connected in an undirected network G where $g_{ij} = 1$ if i and j are linked. Individual i has degree $d_i = \sum_j g_{ij}$, and $P(d)$ denotes the degree distribution. Each person chooses whether to engage, $e_i \in 0, 1$, interpreted as taking a visible action to interact around mental-health or financial concerns (e.g., participating in a savings group or training as a listening volunteer).

Choosing $e_i = 1$ yields a private benefit $b_i > 0$, drawn from an atomless distribution H. Engagement involves a cost $c(d_i, 1 - E[e_{j:j\in N}])$ that increases with both the expected share of disengagement in society $(1 - E[e_{j:j\in N}])$ and the individual's degree (d_i). We assume that c is weakly increasing in both arguments, $\frac{\partial c}{\partial d_i (1-E[e_{j:j\in N})} > 0$, and $c(:, 0) = 0$.

An individual i's utility can be written as:

$$U(e_i, d_i; e_{j:j\in N_i}) = [b_i - c(d_i, 1 - E[e_{j:j\in N}])]e_i \tag{1}$$

Individual i engages if $b_i > c(d_i, 1 - E[e_{j:j \in N}])$, so the engagement probability is $1 - H(c(d_i, 1 - E[e_{j:j \in N}]))$. Agents know their degree, $P(d)$, and H.

7.1 Empirical Evidence for Model Assumptions

A key assumption is that engagement is more costly for more connected individuals when the perceived norm is non-engagement.

Several intuitive mechanisms support this. First, having more connections implies having a larger "audience," raising the reputational risks of sharing sensitive information since more peers can gossip, judge, or update their beliefs negatively (as noted in Section 6). Second, highly connected individuals face stronger pressures to conform to existing norms and have more relationships at stake. In fact, as Young (2015) notes:

> *"Norm entrepreneurship is a risky undertaking for the well-connected. For this reason, politicians, religious leaders, and village elders may be among the least willing to induce norm shifts even though they are the ones most capable of doing it."*

Third, people with many connections often have a combination of strong and weak ties, increasing the probability of encountering someone who might be judgmental or insensitive. Finally, they may simply have less time and attention to devote to such interactions. Our network prediction experiment provides direct support for this: 42% of respondents believe that advice-taking links will not form when the advisor is highly connected because "the advisor will not have time". Finally, we find correlational evidence consistent with this assumption in our data: individuals with more connections are less likely to want to participate in savings groups (Figure G.1).

These mechanisms are further supported by heterogeneity in treatment effects: Consistent with this assumption, more connected individuals respond more strongly to the belief-correction intervention, consistent with a higher marginal cost of engaging when norms are perceived as unfavorable. Regressing the decision to participate in savings groups and listening services on the interaction of degree centrality and beliefs about non-engagement yields negative coefficients, suggesting that pessimistic beliefs deter engagement particularly strongly among the better connected.[26]

7.2 Equilibrium

Following Jackson and Yariv (2007), a symmetric Bayesian Nash equilibrium satisfies:

$$\hat{e} = \sum_d P(d)(1 - H(c(d, 1 - \hat{e}))). \tag{2}$$

We then consider the case, following Jackson (2019), in which individuals naively estimate the social norm using only the behavior of their neighbors. Because of the friendship paradox

[26] The coefficient is also negative when we either use beliefs about others' willingness to engage around financial well-being as a measure of beliefs or when we use savings groups as the measure of engagement

(Feld, 1991), the degree distribution of neighbors is: $\widetilde{P}(d) = \frac{d}{E[d]}P(d)$, which first-order stochastically dominates $P(d)$. Individuals, therefore, infer the norm from a more connected subpopulation, which—because c increases in degree—leads them to overestimate non-engagement.

The perceived (peer-based) equilibrium $\tilde{e}$ satisfies:

$$\widetilde{e} = \sum_d \widetilde{P}(d)(1 - H(c(d, 1 - \widetilde{e}))) = 1 - \sum_d \widetilde{P}(d)(H(c(d, 1 - \widetilde{e}))) \tag{3}$$

where $\tilde{e}$ denotes the equilibrium engagement when individuals use $\widetilde{P}(d)$ instead of $P(d)$. Rewriting the equation succinctly, denoting the probability that a randomly chosen peer does not engage by a, and $\widetilde{a} + \widetilde{e} = 1$:

$$\widetilde{e} = 1 - \sum_d \widetilde{P}(d)(H(c(d, 1 - \widetilde{e}))) = 1 - \widetilde{a} \implies \widetilde{a} = \sum_d \widetilde{P}(d)(H(c(d, \widetilde{a})))$$

Lemma 1 implies that individuals always perceive higher non-engagement when relying on peers to make inferences.[27]:

Lemma 1 $E[H(c(d,a)] \leq \widetilde{E}[H(c(d,a)] \; \forall a \in [0,1]$

The following proposition compares equilibria with and without the friendship paradox in the simple case where there exist multiple (three) equilibria; one is a stable equilibrium at zero and two stable and unstable, strictly positive equilibria – as shown in Figure 4. This follows from standard assumptions about the shape of the CDF H of the cost function that we assume for now, but later estimate empirically.

Proposition 1 *Let $E[a_{j:j \in N}]$ be the society-wide expected non-engagement and a_h and a_l be the non-zero symmetric Bayes Nash equilibrium non-engagement that solves Equation 2 such that $a_h > a_l$. Let $E[a_{j:j \in N(i)}]$ be the expected non-engagement among the agent's friends and let $\widetilde{a}_l$ and $\widetilde{a}_h$ be the symmetric Bayes Nash equilibria that solve Equation 3 such that $\widetilde{a}_l < \widetilde{a}_h$. Then, $\widetilde{a}_l \leq a_l$ and $\widetilde{a}_h \geq a_h$.*

The proof is provided in Appendix Section G.3. Figure 4 illustrates this result. The proposition implies that in a society where there are three possible equilibrium values of non-engagement out of which two are non-zero, one stable, and one unstable, the highest stable equilibrium non-engagement is even higher when individuals observe the actions of their peers and not the entire society. The lower unstable equilibrium is even lower, implying that a larger shift in expected engagement is required to cross the tipping point and reach the zero equilibrium non-engagement.

7.3 Dynamics and the Effect of Belief Correction

Consider the dynamic form of the best response map of Equation 3:[28]

[27] The proof for Lemma 1 is provided in the appendix in Section G.2.

[28] Previous period's non-engagement a_{t-1} is the relevant 'belief' that individuals have about others' willingness to not engage while deciding whether or not to engage in period t.

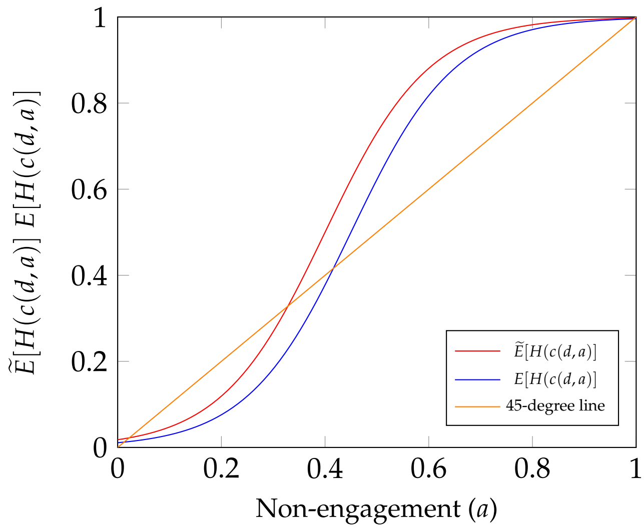


Figure 4: Illustration for Proposition 1 under standard assumptions about H.

$$\widetilde{a}_t = \sum_d \widetilde{P}(d)(H(c(d, \widetilde{a}_{t-1})))$$

Information about the true share of those willing to engage shifts beliefs about $\widetilde{a}_t$, which then affects engagement at $t+1$. This is useful in the case where individuals overestimate non-engagement because they naively proxy for the norm using the actions taken by their peers. Because individuals continue to infer norms from peers, the updated beliefs propagate recursively. The long-run effect depends on whether the information shock pushes the system past the unstable middle equilibrium. If not, the system gradually reverts to the original low-engagement equilibrium; if so, it converges to a higher-engagement one.

In structural estimation, we use the random variation induced by the RCT to recover the shape of H, identify the location of the unstable equilibrium, and evaluate whether the intervention could shift the system to a new long-run equilibrium.

8 Structural Estimation

We structurally estimate the model using the RCT to (i) recover the equilibrium level of engagement in these communities and assess the scope for belief correction, and (ii) compare belief-shifting to alternative policy levers that do not target beliefs.

8.1 Assumptions

We make a few simplifying assumptions. First, the cost of engagement is $c(d, 1 - E[e_{j:j \in N(i)}]) = \theta * d * E[a_{j:j \in N(i)}]$ where $a_i = 1 - e_i$ is the choice to disengage. The benefit of engagement b_i is drawn from a log-logistic distribution H, implying $H(c(d, a)) = (1 + (\theta d a / \alpha)^{-\kappa})^{-1}$ where we set $\alpha = 1$. This functional form allows for complete flexibility in the number and type of possible equilibria.[29] Figure H.1 plots the log-logistic function for different values of θ and κ to illustrate that we do not impose a specific shape.

Based on the evidence of a positive relationship between optimistic beliefs about peers' willingness to engage and own engagement detected from the experiment results, we restrict the strategic complementarity parameter θ to be strictly positive. We restrict κ to be non-negative as the log-logistic distribution only allows a non-negative domain. The estimation recovers estimates of the two parameters that govern equilibrium selection: θ (complementarities in disengagement) and κ (shape of the benefit distribution).

8.2 Estimation Strategy

We use simulated method of moments and proceed in the following steps.

1. Computation of Degree Distribution: The degree distribution $P(d)$ is computed using the overall network degree.[30] The peer degree distribution is: $\widetilde{P}(d) = P(d) * (d/E(d))$, where the sample average degree is an estimate of the expected degree $E(d)$.[31]

2. Equilibrium selection for the control group: For randomly picked initial θ and κ, we solve for the Bayes-Nash equilibria (Equation 3): $a^*(\theta, \kappa) = \sum \widetilde{P}(d) H(c(d, a^*(\theta, \kappa))$. Several equilibria may exist. Following Bisin et al. (2011) and De Paula (2013), we select the equilibrium that maximizes the likelihood of observed disengagement in the control group. We use a_c^* and d_i to compute the expected disengagement for each agent i in the control group. This is used to construct the first moment of interest: expected disengagement in the control group.

3. Treatment group prediction: We then compute the probability that each individual in the treatment group disengages. This probability depends on the exogenously delivered belief that $a_t^* = 0.4$, the individual's degree d_i, and the values of θ and κ. [32] We compute the expected disengagement in the treatment group serves as the second moment of interest.

4. Objective function: We minimize the squared deviation between model-predicted and empirical means of disengagement in the treatment and control groups, using pattern search with many starting points to avoid local minima.[33]

We use willingness to participate in the listening service as the primary engagement measure, since this most closely corresponds to the information delivered. Robustness to alternative measures is presented in Section 8.3.1.

[29] We use this distribution since its cumulative distribution function can be written in closed form, and this helps with accuracy and speed in our simulations. This is also a more suitable choice than the logistic distribution since it has a positive domain and satisfies $H(0) = 0$. The parameters of the logistic distribution only allow $H(0) = 0$ in the limiting case, and we show robustness to this in the Appendix.

[30] Recall that the overall network contains a link between two agents if they interact in any capacity.

[31] We do not compute degree distributions for each NGO center due to small samples in each center. There is no reason to believe that the networks should be systematically different across centers.

[32] We do not have to solve for the equilibrium and pick the one that maximizes the likelihood in this case, since we know that treated individuals were provided with the expected disengagement.

[33] We use the *pattern search* solver with multiple starting points to run these iterations and find the optimal θ and κ. This allows us to find the global optimum (Audet and Dennis Jr, 2002).

8.3 Estimation Results

Table 1 reports the minimized objective function and shows that the estimated parameters fit the data closely. The model also predicts well the non-targeted moments— namely, the standard deviation and skewness of engagement—indicating good overall fit. Figure 5 plots the equilibrium equation $a = \sum_d \widetilde{P}(d)(H(c(d,a);\theta,\kappa))$ evaluated at the estimated θ^* and κ^*.

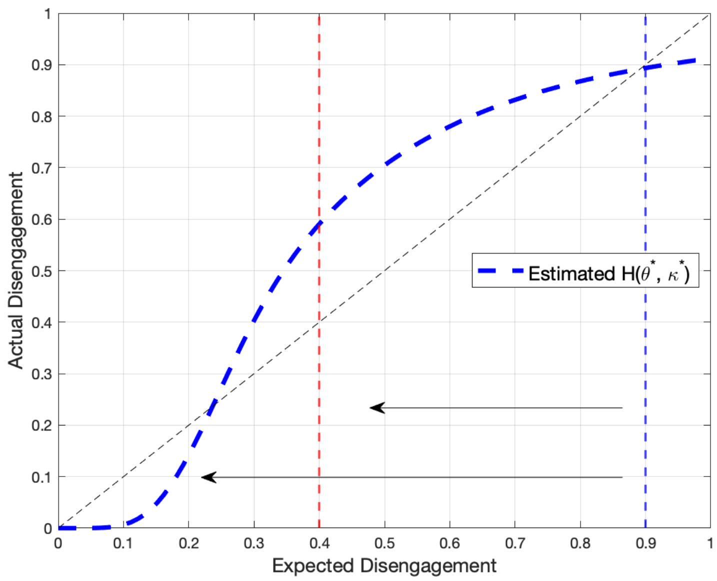
Table 1: Structural Estimation Model Fit

	Targetted Moments				Untargetted Moments	
Objective Function	Predicted Mean (Treated)	Actual Mean (Treated)	Predicted Mean (Control)	Actual Mean (Control)	Standard Deviation (Squared Gap)	Skewness (Squared Gap)
0.0005	0.309	0.299	0.440	0.460	0.003	2.6e-05

Notes: This table shows the model fit for the estimated values of θ^* and κ^* for the main measure of engagement. The objective function is computed at the estimated parameters and is equal to the sum of the squared gap between the mean of the treatment and control groups predicted by the model and in the data respectively.

The estimated best response curve implies three equilibria: two stable equilibria at $a^* = 0$ and $a^* = 0.89$, and an unstable equilibrium at $a^* = 0.23$. We find that the high-disengagement equilibrium at $a^* = 0.89$ maximizes the likelihood, consistent with the low baseline levels of dialogue around these issues. Together, these results indicate that these communities are located near a high-disengagement steady state and that beliefs about disengagement must shift past below the 0.23 threshold to generate persistent change.

Figure 5: Structural Estimation: Actual and Expected Disengagement



Notes: The figure plots actual disengagement a and expected disengagement $\sum_d \widetilde{P}(d)(H(c(d,a)))$ at the estimated θ^* and κ^*. The blue dotted line shows the equilibrium that maximises the likelihood while the red dotted line shows the predicted short-run effect of the intervention.

The structural estimates allow us to assess whether the belief shift induced by the experiment is large enough to generate persistent change. Shifting perceived disengagement to 0.4—as in the RCT—does not cross the unstable equilibrium at $a^* = 0.23$. Because the community sits near the high-disengagement steady state, long-run change would require pushing perceived disengagement to lower than 20% (i.e., perceived engagement to be greater than 80%). Such a shift is substantial and may be difficult to deliver credibly in the status quo, where these topics are not openly discussed. Consistent with this prediction, two years later, individuals exposed to the intervention still hold significantly more optimistic beliefs than others, but their beliefs remain below the level conveyed in the experiment.

8.3.1 Robustness to other measures of engagement and distributional assumptions

We assess robustness using two alternative measures of engagement: (1) willingness to join a savings group and (2) a composite 'community engagement' index (described earlier). Although the treatment information concerned willingness to engage with mental health topics, these outcomes are still plausible proxies for engagement; nevertheless, participation in the listening service remains the closest measure. For both alternative outcomes, we obtain a similar S-shaped best-response curve with three equilibria and a comparable belief threshold (Appendix Table H.1). This implies that sufficiently strong belief shifts would generate long-run change across all measures.

We also examine robustness to the assumed benefit distribution by replacing the log-logistic with a logistic distribution. Re-estimating the model now requires three parameters—θ, the mean κ, and the standard deviation σ—with σ identified using the empirical standard deviation of engagement. As shown in Figure H.2 in the appendix, the resulting best-response curve and implied equilibria are nearly identical to those under the log-logistic specification. We therefore proceed to the counterfactual analysis next.

8.3.2 Counterfactual 1: Increasing the perceived benefits of interactions.

We first consider policies that increase the perceived benefits of engagement—such as information or awareness sessions highlighting the value of network interactions, or financial incentives facilitating interactions. The top panel of Figure 6 plots the estimated best response curve evaluated at the estimated θ^* and κ^* and the dotted line in orange shows how much the curve must shift downward for the long-run effect to be at least as large as the short-run effect of the belief correction intervention.

Formally, shifting the curve is equivalent to increasing the benefits of engagement by a constant $b > 0$. As b rises, expected disengagement falls. We increase b until the resulting long-run equilibrium disengagement is no higher than 0.4, matching the short-run effect of the belief intervention. We find that the mean benefit must increase by roughly 48% to generate a persistent long-run impact of a comparable magnitude, which is a substantially demanding increase. Pushing benefits even higher would collapse disengagement to zero.

8.3.3 Counterfactual 2: Reducing how much agents care about the social norm.

A second class of policies could lower θ, the extent to which violating social norms raises the costs of engagement. Such policies could include psychological interventions that reduce concerns about gossip or judgment, or institutional changes—such as a formal job referral service—that lessen the signaling consequences of seeking help.

We therefore ask: by how much must θ fall to generate a long-run reduction in disengagement at least as large as the RCT's short-run effect? The bottom panel of Figure 6 shows the resulting counterfactual. We gradually decrease θ and compute the implied equilibrium until disengagement falls to 0.4 or lower. We find that θ must fall by roughly 33% relative to the estimated θ^*. This is a sizable reduction, implying that policies aimed at lowering norm-related costs would need to be substantial—and likely costly—to match the short-run impact of belief correction.

Summary of Counterfactual Analysis: Counterfactual analysis shows that generating long-run effects through policies that shift the benefits or costs of engagement can be very demanding. By comparison, belief correction produces sizable short-run effects at low cost and can help increase support for more resource-intensive policies.

9 Conclusion

Social networks cannot serve as effective safety nets if individuals do not demand support. While this concern is relevant across a variety of settings ranging from risk-sharing in communities to advice-taking in organizations, it is especially consequential in environments where individuals lack formal assistance. In such an environment, we show that inaccurate beliefs about others' willingness to engage around financial and mental-health concerns meaningfully hinder useful interactions, generating "silent networks" with limited interaction despite high potential benefits.

Using a field experiment in urban informal communities in India, we demonstrate that correcting these misperceptions increases the demand for network-based support by reducing perceived reputational and interaction costs. Individuals with pessimistic beliefs place the highest weight on these social costs while deciding whether an interaction would take place. Additional experiments confirm these mechanisms and help rule out alternative explanations, including social desirability concerns.

To understand why misperceptions arise, we adapt a network model in which individuals infer the norm from their peers' behavior. Because people's peers are more connected—and therefore face higher social costs of engagement—than a randomly chosen community member, individuals systematically under-estimate others' willingness to interact. This mechanism, rooted in the friendship paradox, can trap communities in a low-interaction equilibrium even when most individuals privately value support.

While structural estimates predict that belief correction alone may be insufficient to move the communities out of a low-engagement equilibrium, counterfactuals indicate that alternative policies for persistent change can be much more demanding. Nevertheless, belief correction can raise the short-run demand for engagement within networks and help generate resources needed to finance more intensive interventions, a valuable feature in underfunded settings.

Future work could examine how individuals update their beliefs about social norms, compare the effectiveness of correcting different types of beliefs for sustaining long-run social interactions, and test the efficacy of providing platforms to encourage social interactions.

Tables and Figures

A Summary Statistics

Table 2: Summary Statistics

	Mean	SD
A. Demographics		
Age	33.75	(9.371)
Female	0.352	(0.478)
Monthly HH Income ($<$ Rs 2,500)	0.157	(0.364)
Monthly HH Income (Rs 2,500-5,000)	0.189	(0.392)
Monthly HH Income (Rs 5,000-10,000)	0.444	(0.498)
Monthly HH Income ($>$ Rs 10,000)	0.210	(0.408)
Employed (Yes/No)	0.668	(0.472)
B. Well-Being		
Stress (Index; Scale 1-5)	3.076	(0.864)
Unable to manage difficulties (Often/Very Often)	0.503	(0.501)
Life Satisfaction (Scale 1-4)	2.851	(0.951)
Happiness (Scale 1-4)	2.479	(0.981)
C. Networks		
Number of network connections (Advice)	2.787	(3.378)
Number of network connections (Overall)	3.954	(3.945)
Observations	352	

Notes: The above table shows the summary statistics (mean and standard deviation) for various demographic characteristics of interest from the baseline sample.

Table 3: Summary Statistics

	Mean	SD
A. Willingness to Engage		
Willingness to Talk (Financial Concerns)	0.707	(0.456)
Willingness to Talk (Mental Health)	0.631	(0.483)
Willingness to Talk (Physical Health)	0.618	(0.487)
Stigma (Physical Health)	0.555	(0.498)
Stigma (Mental Health)	0.503	(0.501)
B. Beliefs		
Beliefs (Mental Health)	4.985	(3.012)
Beliefs (Physical Health)	5.126	(3.029)
Beliefs (Financial Concerns)	5.960	(3.196)
Beliefs -Stigma (Mental Health)	4.802	(3.312)
Beliefs -Stigma (Physical Health)	4.812	(3.253)
Beliefs -Stress (Mental Health)	5.641	(3.313)
C. Frequency of Conversations		
Dialogue (Physical Health; 1-5)	2.180	(1.300)
Physical Health talk (Never)	0.430	(0.496)
Physical Health talk (Rarely)	0.223	(0.417)
Physical Health talk (Sometimes)	0.162	(0.369)
Physical Health talk (Often)	0.110	(0.313)
Physical Health talk (Very Often)	0.0762	(0.266)
Dialogue (Mental Health; 1-5)	2.511	(1.304)
Mental Health talk (Never)	0.292	(0.456)
Mental Health talk (Rarely)	0.252	(0.435)
Mental Health talk (Sometimes)	0.191	(0.394)
Mental Health talk (Often)	0.182	(0.386)
Mental Health talk (Very Often)	0.0831	(0.276)
Dialogue (Financial Concerns; 1-5)	2.847	(1.470)
Financial Concerns talk (Never)	0.256	(0.437)
Financial Concerns talk (Rarely)	0.200	(0.401)
Financial Concerns talk (Sometimes)	0.181	(0.386)
Financial Concerns talk (Often)	0.166	(0.372)
Financial Concerns talk (Very Often)	0.197	(0.398)
Observations	334	

Notes: The above table shows summary statistics (mean and standard deviation) for additional baseline variables such as willingness to engage, dialogue intensity, stigma, and beliefs about peers. All variables in Panel A are binary. We measure stigma by asking if individuals think people should "stay away" from those with specific health concerns. Beliefs about peers are measured in terms of 0-10 individuals in the community. Dialogue intensity is measured on a scale from 1-5 where 1 is 'Never' and 5 is 'Very Often'.

B Endline Results

Table 4: Treatment Effect on Endline Outcomes

Panel A: Treatment Effects on Community Engagement

Variables	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Community Engagement
Treatment	0.147**	0.161**	0.122**	6.577**	0.138***
	(0.0711)	(0.0730)	(0.0612)	(2.959)	(0.0516)
Bootstrap p-value	0.0831	0.0511	0.0561	0.0931	0.0340
q-values	0.046	0.046	0.046	0.046	0.046
Constant	0.667***	0.540***	0.738***	22.72***	0.662***
	(0.0548)	(0.0537)	(0.0483)	(2.090)	(0.0393)
Observations	150	174	170	163	150
R-squared	0.028	0.027	0.023	0.030	0.046

Panel B: Treatment Effects on Own Health Outcomes

	Speaking to the Doctor (MH)	Depression Scoring (Immediate)	Depression Scoring	Listening to Helpline Numbers
Treatment	0.0995*	-0.272***	-0.136*	-0.197**
	(0.0600)	(0.0809)	(0.0747)	(0.0844)
Bootstrap p-value	0.232	0.0300	0.469	0.0511
q-values	0.164	0.008	0.224	
Constant	0.118***	0.606***	0.789***	0.591***
	(0.0373)	(0.0584)	(0.0488)	(0.0610)
Observations	154	143	143	137
R-squared	0.018	0.074	0.023	0.039

Panel C: Treatment Effects on Other Endline Outcomes

VARIABLES	Memory (Numbers Remembered)	Physical Health Dialogue (with Family)
Treatment	-0.271	0.208***
	(0.205)	(0.0711)
Bootstrap p-value	0.387	0.0210
Constant	1.053***	0.184***
	(0.151)	(0.0448)
Observations	153	155
R-squared	0.011	0.053

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Wild bootstrap *p*-values reported (reps=999) using the method outlined in Cameron et al. (2008) where we treat the NGO centre as the cluster unit.

C Longer Term Effects

Table 5: Effects on Network Interactions after 2 weeks

	Financial Concerns			Mental Health Concerns		
	Reached Out (Yes/No)	Number of Friends Reached Out to	Friend Reached Out (Yes/No)	Reached Out (Yes/No)	Number of Friends Reached Out to	Friend Reached Out (Yes/No)
Treatment	0.188**	1.212***	0.146	0.113	0.911**	0.0414
	(0.0930)	(0.427)	(0.0914)	(0.0863)	(0.386)	(0.0833)
q-values	0.052	0.034	0.054	0.034	0.008	0.139
Constant	0.321***	0.877***	0.283***	0.226***	0.500***	0.226***
	(0.0647)	(0.206)	(0.0625)	(0.0580)	(0.158)	(0.0580)
Observations	110	109	109	109	109	109
R-squared	0.036	0.068	0.023	0.016	0.048	0.002

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The q-values (Benjamini et al., 2006) reported in each table treat all the outcomes in this table as multiple hypotheses being tested together. Standard errors are robust.

Table 6: Effects on Beliefs and Willingness to Engage after 2 years

Panel A: Effects on Beliefs about Peers after 2 years

VARIABLES	Beliefs (MH)	Beliefs (MH)	Beliefs (MH)	Beliefs (FC)	Beliefs (FC)	Beliefs (FC)
Previous Participant	0.193			0.00819		
	(0.299)			(0.261)		
Heard Information		0.899***			0.751***	
		(0.324)			(0.259)	
Heard Information and was previous participant			1.442***			0.754**
			(0.387)			(0.307)
Constant	3.607***	3.560***	3.558***	3.658***	3.600***	3.621***
	(0.0869)	(0.0866)	(0.0850)	(0.0930)	(0.0915)	(0.0898)
Observations	467	467	467	467	467	467

Panel B: Effects on Willingness to Engage after 2 years

VARIABLES	Willingness to Engage (MH)	Willingness to Engage (MH)	Willingness to Engage (MH)	Willingness to Engage (FC)	Willingness to Engage (FC)	Willingness to Engage (FC)
Previous Participant	0.0185			0.0611*		
	(0.0448)			(0.0362)		
Heard Information		0.0561			0.0715*	
		(0.0478)			(0.0405)	
Heard Information and was previous participant			0.140***			0.126***
			(0.0165)			(0.0158)
Constant	0.865***	0.863***	0.860***	0.872***	0.874***	0.874***
	(0.0169)	(0.0166)	(0.0165)	(0.0165)	(0.0160)	(0.0158)
Observations	467	467	467	467	467	467

Note: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. We implement post double selection lasso (Belloni et al., 2014) with robust standard errors accounting for income, age, gender, and number of network connections. Previous participant is a binary variable equal to 1 if the respondent in 2023 reports being contacted for the previous study in 2020-21. Heard information is equal to 1 if they report having heard anything about their community's willingness to engage from their peers and 0 otherwise. "MH" refers to mental health concerns and "FC" refers to financial concerns.

Table 7: Effects on Network Interactions and Consumption Outcomes after 2 years

Panel A: Effects on Network Interactions after 2 years

VARIABLES	(1) MH Talk (Above Median)	(2) MH Talk (Above Median)	(3) MH Talk (Above Median)	(4) FC Talk (Above Median)	(5) FC Talk (Above Median)	(6) FC Talk (Above Median)
Previous Participant	0.123*			0.0814		
	(0.0689)			(0.0685)		
Heard Information		0.183**			0.263***	
		(0.0841)			(0.0830)	
Heard Information and was previous participant			0.210**			0.171
			(0.102)			(0.104)
Constant	0.410***	0.412***	0.415***	0.369***	0.358***	0.370***
	(0.0244)	(0.0237)	(0.0234)	(0.0239)	(0.0231)	(0.0229)
Observations	467	467	467	467	467	467

Panel B: Effects on Consumption Outcomes after 2 years

VARIABLES	Consumption Crisis (Yes/No)	Consumption Crisis (Yes/No)	Consumption Crisis (Yes/No)	Consumption varies a lot	Consumption varies a lot
Previous Participant	0.0124			-0.0649**	
	(0.0690)			(0.0275)	
Heard Information		-0.157*			-0.0683**
		(0.0833)			(0.0302)
Heard Information and was previous participant			-0.155		
			(0.102)		
Constant	0.521***	0.535***	0.530***	0.0983***	0.0953***
	(0.0248)	(0.0241)	(0.0237)	(0.0148)	(0.0142)
Observations	467	467	467	467	467

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses. We implement post double selection lasso (Belloni et al., 2014) with robust standard errors accounting for income, age, gender, and number of network connections. *Previous participant* is a binary variable equal to 1 if the respondent in 2023 reports being contacted for the previous study in 2020-21. *Heard information* is equal to 1 if they report having heard anything about their community's willingness to engage from their peers and 0 otherwise. "MH" refers to mental health, and "FC" refers to financial concerns.

D Mechanisms

Table 8: Hypothetical Network Prediction Experiment (OLS)

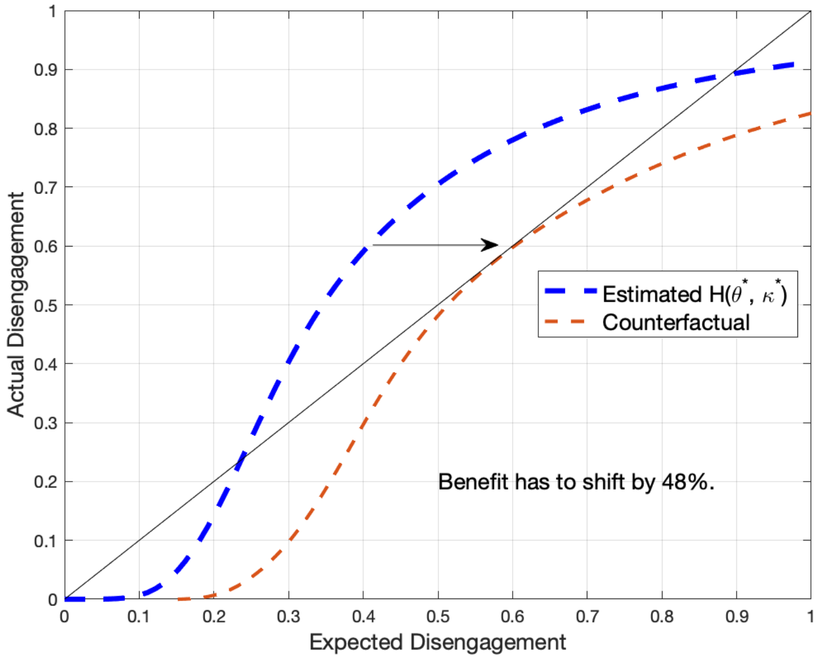
VARIABLES	(1) B takes advice (MH)	(2) B takes advice (MH)	(3) B takes advice (FC)	(4) B takes advice (FC)
Advisor has job contacts	0.00851	0.00687	0.0249	0.0240
	(0.0174)	(0.0173)	(0.0177)	(0.0176)
Advisor is network central	0.0100	0.00740	0.0222	0.0233
	(0.0174)	(0.0173)	(0.0177)	(0.0176)
Advisor has attended training	-0.0181	-0.0182	-0.00929	-0.00972
	(0.0174)	(0.0173)	(0.0177)	(0.0176)
Advisor has job contacts X Own Beliefs		0.0226		0.0158
		(0.0183)		(0.0173)
Advisor is network central X Own Beliefs		0.0347*		0.0354**
		(0.0184)		(0.0173)
Advisor has attended training X Own Beliefs		0.0312*		0.0482***
		(0.0183)		(0.0173)
Beliefs (MH)		-0.00548		
		(0.0184)		
Beliefs (FC)				-0.0635***
				(0.0176)
Constant	0.766***	0.767***	0.737***	0.737***
	(0.0173)	(0.0172)	(0.0178)	(0.0176)
Observations	2,372	2,372	2,370	2,370
R-squared	0.001	0.013	0.002	0.008

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. This table presents the results the link prediction experiment where individuals are asked to predict if a link exists between a hypothetical advisor (A) and a randomly chosen person (B). "MH" refers to mental health and "FC" refers to financial concerns. "Beliefs" refers to the respondent's prediction (standardized) about how many community members would be willing to engage around mental or financial concerns. The interaction terms interact the advisor characteristics with Beliefs (MH) in Column 2 and Beliefs (FC) in Column 4. We report robust standard errors.

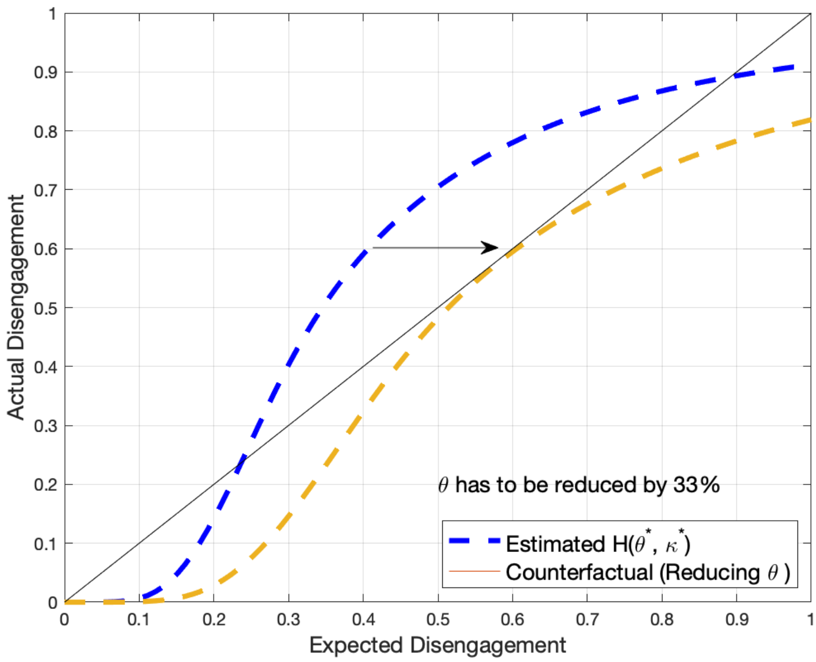
E Structural Estimation: Counterfactual Results

Figure 6: Counterfactual Interventions

(A) Increasing the Benefit of Engagement.



(B) Reducing how much agents care about the Social Norm.



Notes: The figures plot actual disengagement a and expected disengagement $\sum_d \widetilde{P}(d)(H(c(d,a)))$ at the estimated θ^* and κ^* in blue. In addition to this, the figures also plot the proposed counterfactuals that increase benefits (top) and reduce θ (bottom) in the dotted line in orange/yellow.

References

Ambrus, A. and Elliott, M. (2021), 'Investments in social ties, risk sharing, and inequality', *The Review of Economic Studies* **88**(4), 1624–1664.

Anderson, M. L. (2008), 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects', *Journal of the American statistical Association* **103**(484), 1481–1495.

Angelucci, M., De Giorgi, G. and Rasul, I. (2018), 'Consumption and investment in resource pooling family networks', *The Economic Journal* **128**(615), 2613–2651.

Audet, C. and Dennis Jr, J. E. (2002), 'Analysis of generalized pattern searches', *SIAM Journal on optimization* **13**(3), 889–903.

Banerjee, A., Breza, E., Chandrasekhar, A. G. and Golub, B. (2024), 'When less is more: Experimental evidence on information delivery during india's demonetisation', *Review of Economic Studies* **91**(4), 1884–1922.

Banerjee, A., Chandrasekhar, A. G., Duflo, E. and Jackson, M. O. (2013), 'The diffusion of microfinance', *Science* **341**(6144), 1236498.

Beaman, L., BenYishay, A., Magruder, J. and Mobarak, A. M. (2021), 'Can network theory-based targeting increase technology adoption?', *American Economic Review* **111**(6), 1918–1943.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014), 'Inference on treatment effects after selection among high-dimensional controls', *Review of Economic Studies* **81**(2), 608–650.

Benjamini, Y., Krieger, A. M. and Yekutieli, D. (2006), 'Adaptive linear step-up procedures that control the false discovery rate', *Biometrika* **93**(3), 491–507.

Bisin, A., Moro, A. and Topa, G. (2011), The empirical content of models with multiple equilibria in economies with social interactions, Technical report, National Bureau of Economic Research.

Breza, E. (2016), *Field experiments, social networks, and development*, Oxford University Press Oxford, UK.

Breza, E., Chandrasekhar, A., Golub, B. and Parvathaneni, A. (2019), 'Networks in economic development', *Oxford Review of Economic Policy* **35**(4), 678–721.

Bursztyn, L., González, A. L. and Yanagizawa-Drott, D. (2020), 'Misperceived social norms: Women working outside the home in saudi arabia', *American economic review* **110**(10), 2997–3029.

Bursztyn, L. and Yang, D. Y. (2022), 'Misperceptions about others', *Annual Review of Economics* **14**, 425–452.

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2008), 'Bootstrap-based improvements for inference with clustered errors', *The review of economics and statistics* **90**(3), 414–427.

Chandrasekhar, A. G., Duflo, E., Kremer, M., Pugliese, J. F., Robinson, J. and Schilbach, F. (2022), Blue spoons: Sparking communication about appropriate technology use, Technical report, National Bureau of Economic Research.

Chandrasekhar, A. G., Golub, B. and Yang, H. (2019), 'Signaling, shame, and silence in social learning', *Available at SSRN: https://ssrn.com/abstract=3261632 or http://dx.doi.org/10.2139/ssrn.3261632* .

Chaturvedi, B., Varma, A., Mukherjee, C., Chaturvedi, R. and Khan, I. (2018), *Wastepickers: Delhi's Forgotten Environmentalists?*, Chintan Environmental Research and Action Group.

De Paula, A. (2013), 'Econometric analysis of games with multiple equilibria', *Annu. Rev. Econ.* **5**(1), 107–131.

Delavande, A. (2023), Expectations in development economics, *in* 'Handbook of Economic Expectations', Elsevier, pp. 261–291.

Delavande, A., Giné, X. and McKenzie, D. (2011), 'Measuring subjective expectations in developing countries: A critical review and new evidence', *Journal of development economics* **94**(2), 151–163.

Epley, N. (2015), *Mindwise: Why we misunderstand what others think, believe, feel, and want*, Vintage.

Epley, N. and Schroeder, J. (2014), 'Mistakenly seeking solitude.', *Journal of Experimental Psychology: General* **143**(5), 1980.

Fafchamps, M. and Gubert, F. (2007), 'The formation of risk sharing networks', *Journal of development Economics* **83**(2), 326–350.

Feld, S. L. (1991), 'Why your friends have more friends than you do', *American Journal of Sociology* **96**(6), 1464–1477.

Haaland, I., Roth, C. and Wohlfart, J. (2020), 'Designing information provision experiments'.

Ivaschenko, O., Rodriguez Alas, C. P., Novikova, M., Romero Robayo, C., Bowen, T. V. and Zhu, L. (2018), *The state of social safety nets 2018*, number 124300, The World Bank.

Jackson, M. O. (2019), 'The friendship paradox and systematic biases in perceptions and social norms', *Journal of Political Economy* **127**(2), 777–818.

Jackson, M. O., Rodriguez-Barraquer, T. and Tan, X. (2012), 'Social capital and social quilts: Network patterns of favor exchange', *The American Economic Review* **102**(5), 1857–1897.

Jackson, M. O. and Yariv, L. (2007), 'Diffusion of behavior and equilibrium properties in network games', *American Economic Review* **97**(2), 92–98.

Karlan, D., Mobius, M., Rosenblat, T. and Szeidl, A. (2009), 'Trust and social collateral', *The Quarterly Journal of Economics* **124**(3), 1307–1361.

Ligon, E., Thomas, J. P. and Worrall, T. (2002), 'Informal insurance arrangements with limited commitment: Theory and evidence from village economies', *The Review of Economic Studies* **69**(1), 209–244.

Marx, B., Stoker, T. and Suri, T. (2013), 'The economics of slums in the developing world', *Journal of Economic perspectives* **27**(4), 187–210.

Medina, M. (2008), 'The informal recycling sector in developing countries: organizing waste pickers to enhance their impact'.

Möbius, M. M. and Rozenblat, T. S. (2016), 'Informal transfers in networks'.

Morten, M. (2019), 'Temporary migration and endogenous risk sharing in village india', *Journal of Political Economy* **127**(1), 1–46.

Munshi, K. and Rosenzweig, M. (2009), Why is mobility in india so low? social insurance, inequality, and growth, Technical report, National Bureau of Economic Research.

Munshi, K. and Rosenzweig, M. (2016), 'Networks and misallocation: Insurance, migration, and the rural-urban wage gap', *The American Economic Review* **106**(1), 46–98.

Perkins, H. W., Haines, M. P. and Rice, R. (2005), 'Misperceiving the college drinking norm and related problems: a nationwide study of exposure to prevention information, perceived norms and student alcohol misuse.', *Journal of studies on alcohol* **66**(4), 470–478.

Perkins, H. W., Meilman, P. W., Leichliter, J. S., Cashin, J. R. and Presley, C. A. (1999), 'Misperceptions of the norms for the frequency of alcohol and other drug use on college campuses', *Journal of American College Health* **47**(6), 253–258.

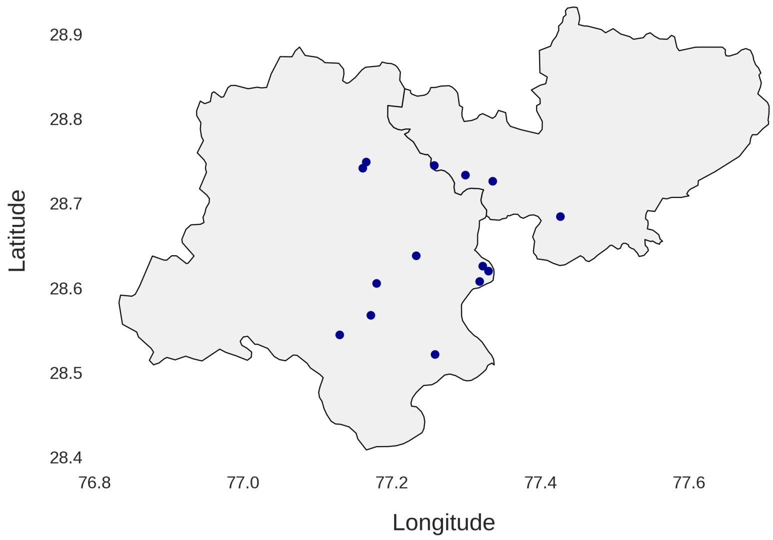
UN-Habitat (2013), *State of the world's cities 2012/2013: Prosperity of cities*, Routledge.

Young, H. P. (2015), 'The evolution of social norms', *economics* **7**(1), 359–387.

Supplementary Appendix

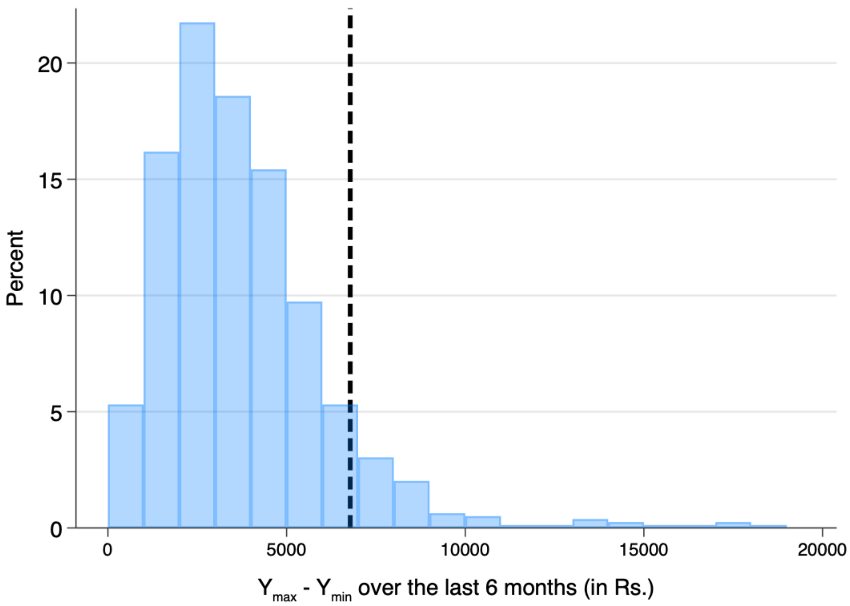
A Baseline

Figure A.1: Survey Locations in the National Capital Region



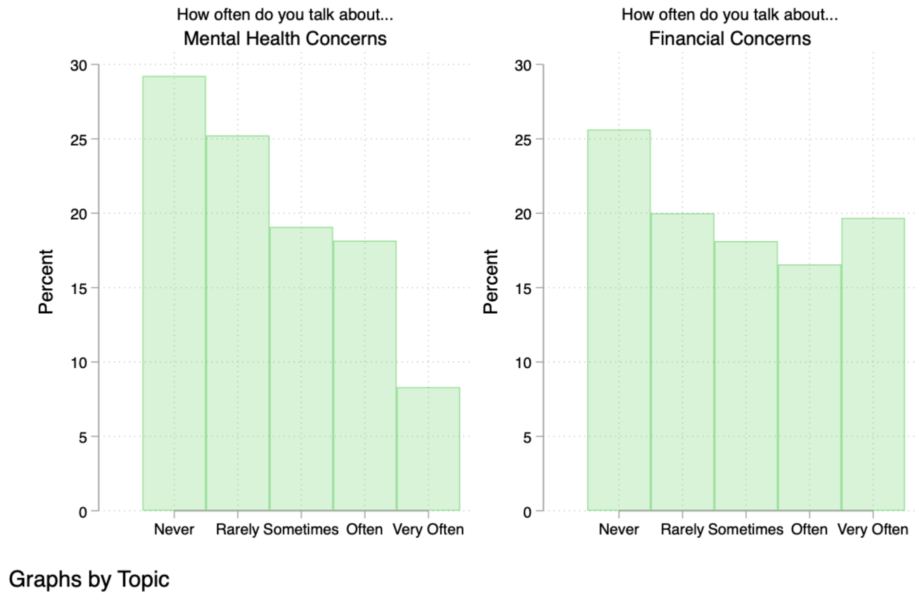
Notes: The figure plots the locations of various centers where we conducted the surveys (in blue) on the map of the national capital region.

Figure A.2: Volatility of Incomes across 6 months



Notes: The figure plots the difference between the maximum and minimum income an individual has earned in the last six months (in Rs). The dashed line is the average current income in the sample. This data was collected for the additional sample in 2023.

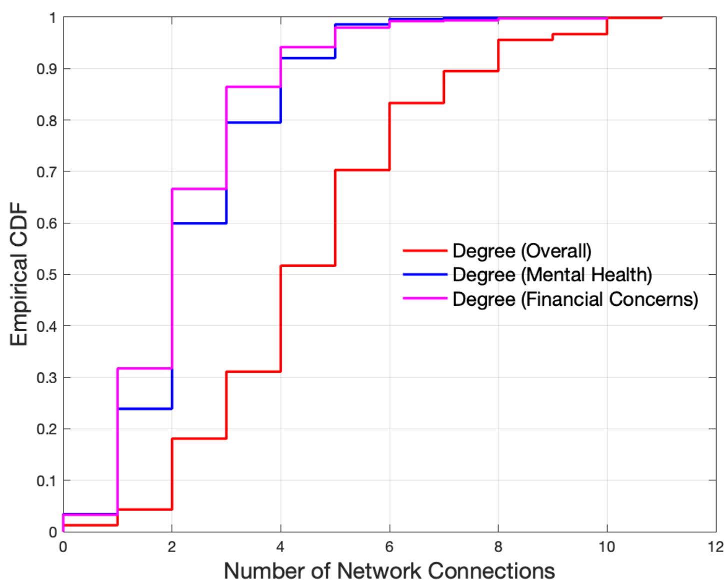
Figure A.3: Intensity of Dialogue in the last two weeks



Notes: The figure plots the baseline intensity of dialogue around mental health and financial concerns with peers in the last two weeks.

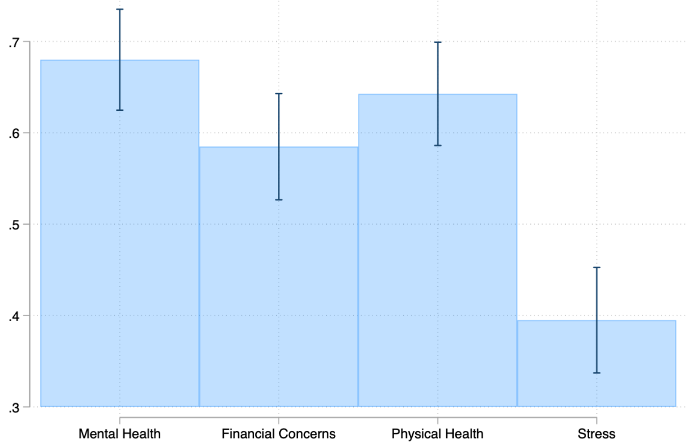
Click to go back.

Figure A.4: Degree Distributions for Overall Networks, Mental Health Advice Networks, and Financial Networks in 2023



Notes: This figure plots the CDF of the number of connections in the overall networks, advice networks, and financial networks for borrowing and lending. Individuals were asked to list up to ten other individuals in their community whom they interact with to borrow/lend, take/give advice, work with etc. ("Overall"), whom they only take advice from regarding mental health issues ("Mental Health"), and whom they contact for financial support ("Financial Concerns").

Figure A.5: Percentage of Underestimators



Notes: This figure plots proportion of underestimators by type of dialogue and additionally by "Stress" where individuals are asked to predict how many others do they think would say that their "difficulties were piling up so high that they could not overcome them" often or very often. Underestimators are identified by comparing their beliefs about community's willingness to engage with actual average in the community (i.e., NGO center).

Table A.1: Correlations among Beliefs and Engagement along various dimensions

	Beliefs (FC)	Beliefs (PH)	Beliefs (MH)	Willing to engage (FC)	Willing to engage (PH)	Willing to engage (MH)
Beliefs (FC)	1					
Beliefs (PH)	0.591***	1				
Beliefs (MH)	0.588***	0.496***	1			
Willing to engage (FC)	0.151*	0.172**	0.206***	1		
Willing to engage (PH)	0.0586	0.0860	0.181**	0.446***	1	
Willing to engage (MH)	0.0516	0.137*	0.106	0.440***	0.550***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents correlations between beliefs and willingness to engage along various topics. The entry in the i'th row and j'th column reports the correlation coefficient between the i'th and j'th variable. "MH" refers to mental health, "PH" refers to physical health, and "FC" refers to financial concerns. Beliefs refer to how many community members out of any 10 individuals believe would be willing to engage around a topic. Willing to engage is a binary variable equal to 1 if they themselves are willing to engage with others around a topic.

Table A.2: Correlations of Network Gaps and Dialogue Intensity

Panel A: Correlations with Beliefs

VARIABLES	Degree Gap	Dialogue (MH)	Dialogue (PH)	Dialogue (FC)
Beliefs (MH)	-0.015	0.064*		
	(0.011)	(0.033)		
Beliefs (PH)			0.021	
			(0.024)	
Beliefs (FC)				0.085***
				(0.021)
Observations	210	271	275	263
R-squared	0.015	0.022	0.002	0.034

Panel B: Correlations with Underestimation

VARIABLES	Degree Gap	Degree Gap	Dialogue (MH)	Dialogue (MH)
Underestimator	-0.015		-0.337*	
	(0.065)		(0.169)	
Severe Underestimator		0.178**		-0.373*
		(0.070)		(0.183)
Observations	210	210	271	271
R-squared	0.000	0.039	0.015	0.015

Note: This table reports regression results where we regress degree gaps and dialogue intensity on (a) various measures of beliefs and (b) whether an individual underestimates engagement with mental health. *Underestimator* is equal to 1 if an individual underestimates how many others in their community are willing to engage. *Severe Underestimator* is equal to 1 if the difference between actual willingness to engage and their belief is greater than the 75th percentile. We construct "Degree Gap" as a network-based measure of an individual's disengagement using differences in degrees in overall and advice networks. "MH" refers to mental health, "PH" refers to physical health, and "FC" refers to financial concerns. Standard errors are robust and clustered at the center level.

Table A.3: Correlations of Willingness to Engage with Beliefs about Community.

VARIABLES	(1) Willing to have MH Dialogue	(2) Willing to have MH Dialogue	(3) Willing to have PH Dialogue	(4) Willing to have PH Dialogue
Beliefs (MH)	0.016			
	(0.012)			
Beliefs -Stigma (MH)		-0.012*		
		(0.006)		
Beliefs (PH)			0.014	
			(0.010)	
Beliefs -Stigma (PH)				-0.023**
				(0.008)
Observations	274	277	275	276
R-squared	0.010	0.007	0.007	0.023

Notes: This table reports results where we regress willingness to engage in dialogue on beliefs about peers willingness to engage and beliefs about stigma among peers. 'MH" refers to mental health, "PH" refers to physical health, and "FC" refers to financial concerns. Standard errors are robust and clustered at the level of the center.

Table A.4: Correlation between Beliefs and Demographic/Network Characteristics (in 2023)

VARIABLES	(1) Belief (MH)	(2) Belief (FC)
Male	-0.109	0.0199
	(0.132)	(0.139)
Age	-0.000661	0.00437
	(0.00599)	(0.00607)
Income	4.22e-05*	1.15e-05
	(2.32e-05)	(2.33e-05)
Degree (Overall)	-0.150***	-0.171***
	(0.0510)	(0.0510)
Degree (MH)	0.265***	0.129*
	(0.0664)	(0.0681)
Degree (FC)	0.351***	0.312***
	(0.0609)	(0.0591)
Talks to peers (MH)	0.0708	0.208
	(0.153)	(0.182)
Talks to peers (FC)	-0.177	-0.0771
	(0.146)	(0.174)
Years in Location	-0.00832	0.00474
	(0.00532)	(0.00583)
Consumption Crisis	0.277*	0.167
	(0.142)	(0.147)
Volatile Consumption	-0.405***	-0.113
	(0.134)	(0.138)
Happiness	0.204***	0.342***
	(0.0608)	(0.0632)
Constant	2.512***	1.798***
	(0.393)	(0.393)
Observations	775	775
R-squared	0.109	0.085

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table presents the regression results where we regress beliefs about community members' willingness to engage around mental health concerns and financial concerns on demographic characteristics, network characteristics, dialogue intensity, and economic indicators. "MH" refers to mental health concerns and "FC" refers to financial concerns. Talks to peers (MH)/ Talks to peers (FC) is a dummy equal to 1 if the individual reports having spoken about mental health or financial concerns to their peers often/very often in the last 2 weeks. Degree refers to number of connections. Happiness is defined on a scale of 1-5. Volatile consumption is a dummy equal to 1 when the individual reports that their consumption has fluctuated a little/a lot over the last 6 months. Consumption crisis is a dummy equal to 1 if the individual reports having faced consumption crisis events very often (i.e. not having enough financial resources to meet basic dietary needs, education expenditures, or health expenditures).

B Design

B.1 Endline Balance

Table B.1: Endline Balance: Demographics and Health

	Control	Treatment	p-value
Age	34.69	34.66	0.98
Female	0.30	0.34	0.60
Income (Category)	2.81	2.83	0.89
Monthly HH Income (< Rs 2,500)	0.14	0.11	0.48
Monthly HH Income (Rs 2,500-5,000)	0.16	0.22	0.30
Monthly HH Income (Rs 5,000-10,000)	0.44	0.41	0.65
Monthly HH Income (Rs 10,000+)	0.26	0.27	0.89
Employed	0.67	0.68	0.91
Stress (Index; Scale 1-4)	3.10	3.07	0.82
Satisfaction (Scale 1-4)	2.84	2.94	0.47
Happiness (Scale 1-4)	2.41	2.48	0.65
Degree (Advice)	2.59	3.14	0.28
Degree (Overall)	3.84	4.57	0.24
Degree (Advice Network) >0	4.67	4.76	0.88
Degree (Overall Network) >0	5.79	5.91	0.85
N	92	88	

Notes: This table presents the results of balance tests between the treatment and control group for the endline sample where the first two columns show the means of the variables for the Control and Treatment Group and the third column shows the *p*-values for the two-sided test that the two means are equal using robust standard errors. *Degree* refers to the number of connections and *Degree* (> 0) only includes strictly positive entries.

Table B.2: Endline Balance: Willingness to Talk and Beliefs

	Control	Treatment	p-value
Willingness to Talk(Financial)	0.70	0.74	0.59
Willingness to Talk (Mental Health)	0.67	0.56	0.12
Willingness to Talk (Physical Health)	0.60	0.63	0.68
Stigma (Physical Health)	0.63	0.53	0.21
Stigma (Mental Health)	0.49	0.50	0.94
Beliefs (Mental Health)	4.70	5.18	0.33
Beliefs (Physical Health)	5.08	5.26	0.73
Beliefs (Financial Concerns)	5.59	6.71	0.03
Beliefs -Stigma (Mental Health)	4.84	5.05	0.70
Beliefs -Stigma (Physical Health)	4.95	5.05	0.84
Beliefs -Stress (Mental Health)	5.67	6.36	0.20
Dialogue (Physical Health; 1-5)	2.22	2.08	0.51
Physical Health talk (Never)	0.45	0.47	0.78
Physical Health talk (Rarely)	0.18	0.19	0.88
Physical Health talk (Sometimes)	0.16	0.19	0.59
Physical Health talk (Often)	0.11	0.07	0.34
Physical Health talk (Very Often)	0.09	0.07	0.64
Dialogue (Mental Health; 1-5)	2.50	2.39	0.58
Mental Health talk (Never)	0.34	0.33	0.87
Mental Health talk (Rarely)	0.17	0.30	0.05
Mental Health talk (Sometimes)	0.21	0.13	0.19
Mental Health talk (Often)	0.21	0.14	0.27
Mental Health talk (Very Often)	0.07	0.10	0.53
Dialogue (Financial Concerns ; 1-5)	2.81	3.01	0.37
Financial Concerns talk (Never)	0.27	0.20	0.28
Financial Concerns talk (Rarely)	0.18	0.21	0.56
Financial Concerns talk (Sometimes)	0.18	0.18	1.00
Financial Concerns talk (Often)	0.20	0.18	0.70
Financial Concerns talk (Very Often)	0.17	0.23	0.33
N	92	88	

Notes: This table presents additional results of balance tests between the treatment and control group for the endline sample where the first two columns show the means of the variables for the Control and Treatment Group and the third column shows the *p*-values for the two-sided test that the two means are equal using robust standard errors. All variables are binary except beliefs and dialogue intensity. Beliefs are measured in terms of 0-10 individuals in the community. Dialogue intensity is measured on a scale from 1-5 where 1 is 'Never' and 5 is 'Very Often'.

C Endline Results

Table C.1: Effects on Additional Endline Outcomes

Panel A: Effects on Demand for Information

	Information Session	Listening to Good Practices (Immediate)	Listening to Good Practices
Treatment	0.0534	-0.0507	0.0263
	(0.0685)	(0.0846)	(0.0636)
Bootstrap p-value	0.430	0.511	0.697
q-values	1	1	1
Constant	0.674***	0.551***	0.821***
	(0.0491)	(0.0603)	(0.0472)
Observations	180	141	139
R-squared	0.003	0.003	0.001

Panel B: Effects on Self Efficacy

VARIABLES	Goals (Finance)	Goals (Education)	Goals (Business)	Self Efficacy
Treatment	0.154	-0.200	0.0436	-0.0268
	(0.264)	(0.243)	(0.278)	(0.222)
Bootstrap p-value	0.647	0.468	0.852	0.893
q-values	1	1	1	1
Constant	2.219***	2.729***	2.548***	2.481***
	(0.176)	(0.183)	(0.190)	(0.151)
Observations	148	140	144	139
R-squared	0.002	0.005	0.000	0.000

Panel C: Effects on Stigma

VARIABLES	Stigma (Information Session)	List Count	Depression Score Revelation
Treatment	0.0313	-0.0763	-0.0917
	(0.0430)	(0.211)	(0.0881)
Bootstrap p-value	0.373	0.781	0.451
q-values	1	1	1
Constant	0.922***	3.507***	0.925***
	(0.0338)	(0.157)	(0.0423)
Observations	128	154	64
R-squared	0.004	0.001	0.020

Robust standard errors in parentheses. Wild bootstrap p-value reported, reps=999.
*** p<0.01, ** p<0.05, * p<0.1

Notes: We report robust standard errors. We also report wild bootstrap p values using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit. The q-values (Benjamini et al., 2006) reported in each table treat the outcomes in the table as multiple hypotheses being tested together.

C.1 Heterogeneity

Table C.2: Treatment Effects by Baseline Willingness to Talk about Mental Health

Panel A: Effect on Engagement Outcomes

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.374***	0.434***	0.204*	11.86**	7.817	0.315***
	(0.137)	(0.119)	(0.119)	(4.964)	(5.063)	(0.104)
Willingness to talk (Mental Health)	0.466***	0.375***	0.186	10.19**	6.190	0.295***
	(0.122)	(0.112)	(0.114)	(4.441)	(4.641)	(0.0949)
Interaction	-0.265*	-0.355**	-0.106	-6.677	-5.553	-0.235**
	(0.154)	(0.151)	(0.138)	(6.223)	(6.075)	(0.117)
Constant	0.350***	0.280***	0.625***	15.87***	26.07***	0.467***
	(0.108)	(0.0909)	(0.100)	(3.656)	(4.083)	(0.0876)
Observations	138	160	156	150	121	138
R-squared	0.185	0.103	0.053	0.069	0.029	0.152
Bootstrap p-value	0.0390	0.101	0.345	0.114	0.187	0.0561

Panel B: Effect on Other Outcomes

VARIABLES	Information Session	Listening to Good Practices (Immediate)	Listening to Good Practices
Treatment	0.444***	0.0292	0.168
	(0.116)	(0.152)	(0.129)
Willingness to talk (Mental Health)	0.488***	0.0389	0.182
	(0.105)	(0.138)	(0.120)
Interaction	-0.510***	-0.144	-0.184
	(0.142)	(0.187)	(0.150)
Constant	0.333***	0.526***	0.684***
	(0.0918)	(0.116)	(0.108)
Observations	164	132	131
R-squared	0.143	0.010	0.027
Bootstrap p-value	0.0120	0.471	0.0340

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows heterogeneous effects on the demand for additional information by baseline willingness to engage with mental health related concerns. We report robust standard errors. We also report wild bootstrap *p*-values for the interaction term using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit.

Click to go back.

Table C.3: Heterogeneous Treatment Effects by Baseline Dialogue

Panel A: Baseline Dialogue Around Mental Health

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.0227	0.0581	0.0808	8.129*	5.643	0.0437
	(0.0948)	(0.107)	(0.0651)	(4.226)	(3.770)	(0.0665)
Baseline MH Dialogue (Below Median)	-0.305***	-0.195*	-0.264***	-12.53***	-6.565*	-0.243***
	(0.107)	(0.106)	(0.0943)	(3.957)	(3.702)	(0.0753)
Interaction	0.282**	0.206	0.119	-0.164	-2.510	0.217**
	(0.142)	(0.146)	(0.117)	(5.632)	(5.189)	(0.100)
Constant	0.805***	0.636***	0.864***	28.90***	33.86***	0.772***
	(0.0627)	(0.0734)	(0.0524)	(2.951)	(2.661)	(0.0458)
Observations	150	174	170	163	129	150
R-squared	0.088	0.048	0.093	0.137	0.078	0.118
Bootstrap p-value	0.0931	0.185	0.296	0.979	0.456	0.108

Panel B: Baseline Dialogue Around Financial Concerns

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.0368	-0.0950	-0.0150	0.887	1.296	0.00919
	(0.0771)	(0.1000)	(0.0847)	(4.532)	(3.980)	(0.0600)
Baseline FC Dialogue (Below Median)	-0.363***	-0.453***	-0.204**	-9.458**	-4.641	-0.317***
	(0.0974)	(0.0951)	(0.0912)	(4.170)	(3.852)	(0.0673)
Interaction	0.183	0.436***	0.237**	9.400	3.410	0.220**
	(0.130)	(0.138)	(0.119)	(5.948)	(5.408)	(0.0937)
Constant	0.875***	0.806***	0.857***	28.44***	33.70***	0.844***
	(0.0593)	(0.0667)	(0.0599)	(3.246)	(2.848)	(0.0422)
Observations	150	174	170	163	129	150
R-squared	0.133	0.133	0.056	0.059	0.022	0.176
Bootstrap p-value	0.154	0.00100	0.0811	0.0280	0.312	0

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows heterogeneous effects on community engagement by a binary variable which indicates whether baseline frequency of dialogue with peers around mental health financial concerns respectively was less than/equal to the median. We report robust standard errors. We also report wild bootstrap p-values for the interaction term using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit.

Table C.4: Heterogeneous Treatment Effects by Degree (Overall)

	(1) Savings Group	(2) Listening Volunteer	(3) Listening Contribution	(4) Contribution (in Rupees)	(5) Contribution (>0)	(6) Community Engagement
Treatment	0.0117	0.159	0.150*	2.992	-3.434	0.117
	(0.110)	(0.107)	(0.0901)	(4.303)	(3.632)	(0.0774)
Degree (Overall)	-0.00736	0.0104	0.00258	-0.464	-0.893*	0.00156
	(0.0137)	(0.0126)	(0.0128)	(0.500)	(0.464)	(0.00846)
Interaction	0.0307*	-0.00116	-0.00625	0.838	1.532**	0.00432
	(0.0183)	(0.0172)	(0.0160)	(0.733)	(0.632)	(0.0123)
Constant	0.696***	0.499***	0.728***	24.59***	34.88***	0.656***
	(0.0755)	(0.0746)	(0.0707)	(3.052)	(2.466)	(0.0572)
Observations	150	174	170	163	129	150
R-squared	0.052	0.034	0.024	0.038	0.050	0.049
Bootstrap p-value	0.0450	0.910	0.566	0.153	0.0651	0.633

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows heterogeneous effects on various measures of community engagement by the number of connections in the overall network. We report robust standard errors. We also report wild bootstrap p-values for the interaction term using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit.

Table C.5: Treatment Effects by Stigma (Mental Health)

	(1) Savings Group	(2) Listening Volunteer	(3) Listening Contribution	(4) Contribution (in Rupees)	(5) Contribution (>0)	(6) Community Engagement
Treatment	0.202**	0.146	0.125	3.224	-0.930	0.149**
	(0.0982)	(0.105)	(0.0866)	(4.171)	(3.726)	(0.0703)
Stigma (Mental Health)	0.110	-0.148	-0.0453	-7.107*	-7.071*	-0.0175
	(0.102)	(0.106)	(0.0884)	(4.262)	(4.033)	(0.0732)
Interaction	-0.0907	0.0921	-0.00473	7.291	9.349*	-0.0119
	(0.144)	(0.150)	(0.124)	(6.001)	(5.470)	(0.103)
Constant	0.632***	0.610***	0.775***	26.25***	33.87***	0.684***
	(0.0685)	(0.0742)	(0.0613)	(2.912)	(2.695)	(0.0491)
Observations	139	161	157	151	122	139
R-squared	0.045	0.053	0.028	0.050	0.040	0.056
Bootstrap p-value	0.527	0.451	0.951	0.274	0.00300	0.854

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: This table shows heterogeneous effects on various measures of community engagement by whether the individual has stigma against mental health during the baseline in that they believe that "people should stay away from those with mental health issues". We report robust standard errors. We also report wild bootstrap p-values for the interaction term using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit.

Click to go back.

Table C.6: Treatment Effects by Underestimators

Panel A: Underestimation of Willingness to Engage around Mental Health

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.0804	0.0833	0.117	11*	8.636	0.110
	(0.116)	(0.130)	(0.127)	(5.723)	(5.587)	(0.0759)
Underestimator (MH)	-0.131	-0.287**	0.0192	5.667	6.329	-0.104
	(0.122)	(0.130)	(0.125)	(5.316)	(5.263)	(0.0820)
Interaction	0.0799	0.128	0.00541	-6.310	-7.626	0.0215
	(0.149)	(0.163)	(0.147)	(6.923)	(6.588)	(0.101)
Constant	0.813***	0.750***	0.750***	19.33***	26.36***	0.771***
	(0.0992)	(0.110)	(0.110)	(4.610)	(4.724)	(0.0650)
Observations	126	146	144	138	113	126
R-squared	0.045	0.091	0.024	0.036	0.024	0.080
Bootstrap p-value	0.632	0.429	0.974	0.402	0.224	0.848

Panel B: Underestimation of Willingness to Engage around Financial Concerns

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.107	0.200	0.100	7.347	5.196	0.118
	(0.122)	(0.135)	(0.114)	(5.352)	(4.983)	(0.0874)
Underestimator (FC)	-0.0295	0.0294	0	1.562	1.786	-0.00128
	(0.126)	(0.134)	(0.117)	(5.032)	(4.698)	(0.0857)
Interaction	0.0387	0.0206	0.0389	-0.694	-2.689	0.0325
	(0.154)	(0.169)	(0.141)	(6.706)	(6.161)	(0.109)
Constant	0.737***	0.500***	0.750***	21.84***	29.64***	0.684***
	(0.103)	(0.113)	(0.0982)	(4.209)	(3.967)	(0.0689)
Observations	126	147	144	138	111	126
R-squared	0.026	0.045	0.024	0.032	0.015	0.053
Bootstrap p-value	0.824	0.838	0.851	0.952	0.705	0.737

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: This table shows heterogeneous effects on community engagement by whether the individual underestimates community willingness to engage around mental health (MH) and financial concerns (FC respectively. We report robust standard errors. We also report wild bootstrap *p*-values for the interaction term using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit.

Click to go back.

Table C.7: Heterogeneous Treatment Effects by Baseline Beliefs about Stress

Panel A: Underestimation of Stress in the Community

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.102	0.192*	0.0901	4.487	1.750	0.110
	(0.0958)	(0.102)	(0.0899)	(4.328)	(3.940)	(0.0706)
Underestimator (Stress)	-0.0210	-0.00905	0.0444	0.590	-1.700	-0.00560
	(0.116)	(0.119)	(0.104)	(4.609)	(4.214)	(0.0814)
Interaction	0.0534	-0.0134	0.0870	5.146	2.677	0.0400
	(0.148)	(0.164)	(0.124)	(6.308)	(5.773)	(0.102)
Constant	0.735***	0.538***	0.737***	23.47***	32.50***	0.696***
	(0.0769)	(0.0809)	(0.0724)	(3.274)	(2.998)	(0.0534)
Observations	128	149	146	140	113	128
R-squared	0.024	0.039	0.035	0.036	0.011	0.046
Bootstrap p-value	0.550	0.940	0.290	0.436	0.607	0.541

Panel B: Beliefs about Stress in the Community

	Savings Group	Listening Volunteer	Listening Contribution	Contribution (in Rupees)	Contribution (>0)	Community Engagement
Treatment	0.143	0.200	0.140	7.454	3.895	0.141
	(0.131)	(0.140)	(0.104)	(5.039)	(4.644)	(0.0891)
Stress Belief (>4)	0.0965	0.0837	-0.0192	2.335	4.387	0.0665
	(0.120)	(0.122)	(0.107)	(4.668)	(4.312)	(0.0839)
Interaction	-0.0393	-0.0248	-0.0380	-2.293	-1.933	-0.0305
	(0.157)	(0.170)	(0.133)	(6.484)	(5.881)	(0.110)
Constant	0.667***	0.481***	0.769***	22.31***	29***	0.653***
	(0.0978)	(0.0975)	(0.0838)	(3.568)	(3.407)	(0.0676)
Observations	128	149	146	140	113	128
R-squared	0.031	0.044	0.023	0.027	0.022	0.051
Bootstrap p-value	0.679	0.871	0.758	0.784	0.770	0.721

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Panel A shows shows heterogeneous effects on various measures of community engagement by whether the individual underestimates the level of stress in their community i.e. their belief is less than the true community-level stress. Panel B shows heterogeneous effects on various measures of community engagement by whether the individual's belief about the level of stress in their community is above the proportion who are actually not willing to engage (i.e. 40%). This is to check if those who think more than 40% are stressed and might assume that these individuals are not willing to engage change their behaviour when they are told that majority are willing to engage. We report robust standard errors. We also report wild bootstrap *p*-values for the interaction term using the method outlined in Cameron et al. (2008) where we treat the NGO center as the cluster unit.

Click to go back.

D Mechanisms

Table D.1: List Experiment to detect Social Desirability Concerns

	3 statements + Statement about engagement with MH	3 statements	Average Difference between 4 and 3 statements
Mean Agreements	2.971	2.217	0.754***
Observations			670

	3 statements + 2 statements about engagement with MH and FC	3 statements	Average Difference between 5 and 3 statements
Mean Agreements	3.717	2.217	1.499***
Observations			504

Notes: These tables present the results of the list experiments where individuals were randomly divided into groups and either asked how many of the three statements they agree with (i.e., control) or asked how many of 4 and 5 statements they agree with (i.e., treated). The additional statements were relating to engagement with mental health and financial concerns respectively. "MH" refers to mental health and "FC" refers to financial concerns. The first two columns show the mean number of agreements when different number of statements are provided. The third column computes the difference between treated and control conditions.

Table D.2: Additional Experiment to detect Social Desirability Concerns

VARIABLES	(1) Savings Group	(2) Listening Service	(3) Savings Group	(4) Listening Service
Increased Distance between Enumerator and Respondent	0.00398	0.0326	0.0241	0.0497
	(0.0280)	(0.0324)	(0.0365)	(0.0453)
Treatment			-0.0609	-0.0305
			(0.0397)	(0.0467)
Increased Distance x Treated			-0.0352	-0.0311
			(0.0555)	(0.0647)
Constant	0.807***	0.690***	0.838***	0.706***
	(0.0199)	(0.0233)	(0.0263)	(0.0326)
Observations	791	791	791	791
R-squared	0.000	0.001	0.011	0.004

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table presents the results of an experiment where we randomly vary whether the enumerator enters the respondent's answer (*Increased Distance*=0) or gives them the device to answer it privately (*Increased Distance*=1). We also interact it with the treatment to see if the response being visible to the enumerator affects responses differently across the two groups. We report robust standard errors.

E Follow up Survey

E.1 Follow up Survey Outcomes

The follow up survey outcomes are listed below.

1. Well-being: Happiness, Life-Satisfaction, Stress
2. Self-reported Dialogue
   (a) Indicator whether they have initiated a conversation with their peers to discuss matters relating to either of their mental and financial concerns.
   (b) Number of peers they have initiated conversations with to discuss matters relating to either of their mental and financial concerns.
   (c) Indicator whether their peers have initiated a conversation with them to discuss matters relating to their own mental and financial concerns.
3. Physical Health Safety
   (a) How often did they follow COVID rules (distancing, wearing masks)?
   (b) How often did they wear safety gear (masks, gloves, jacket) while working in hazardous conditions?
4. Altruism towards community members (Dictator Game)
5. Self-reported contact with a doctor/helpline for mental or physical health concerns

E.2 Follow up Survey: Main Results

The results on network interactions have been presented in the main text in Table 5. Next, we find in Table E.1 that individuals in the treatment group are likely to donate Rs. 31 more than the control group in a hypothetical dictator game where they can split Rs. 200 with a randomly chosen member of their community. This result is robust to the inclusion of unbalanced controls as well. However, we also document that individuals in the treatment group are more likely to report higher levels of unhappiness, stress, and lower life satisfaction. At the same time, we also find that the treatment group wears work safety gear significantly more often than the control group, conditional on having worked. Given that we find an (a) increase in dialogue around mental health and (b) investments in self-care (in terms of greater adoption of safety gear), a potential mechanism explaining lower well-being could be that the role of dialogue in alleviating stress becomes ineffective in the presence of correlated risks such as COVID-19.

However, in the absence of real-time data on the location of these individuals, we acknowledge that we cannot rule out the possibilities that (a) the treatment group has lower well-being because they have migrated to an area of low COVID risk and hear about the impact of the pandemic on their peers via increased dialogue or (b) the treatment group is located in an area of high COVID risk and well being is consequently lower. We have checked balance on several baseline variables and do not ex-ante see any reason why the treatment group will be more or less likely to migrate. However, if this were to be the case, we believe that mechanism (a) seems to be more likely. This is because the result in Tables E.1 shows that the treatment group wears COVID-19 masks less often than the control group. 92% of our sample reports to have *always* worn a mask in the past two weeks before the call and the negative effect on the treatment only driven by the remaining 8%. Individuals in low risk-areas may have been less likely to wear COVID-19 safety masks. Additionally, the proportion of individuals for whom the question on safety gear adoption is missing, is higher for the treatment group after controlling for unbalanced baseline variables. This may imply that they are away from their place of work.

Thus, while we document positive effects on dialogue, altruism, and safety gear adoption, we also document negative effects on happiness, life-satisfaction, and stress. These results highlight that the correlated nature of risk can affect the success of belief-shifting interventions.

Table E.1: Other Outcomes

Panel A: Effects on Other Outcomes

VARIABLES	Masks (How often)	Safety Gear (How often)	Maintain Distance (How often)	Mental Health-Doctor (Made Call)
Treatment	-0.179	0.783***	-0.0351	-0.0229
	(0.108)	(0.223)	(0.145)	(0.0534)
q-values	0.024	0.008	0.19	0.14
Constant	3.943***	2.977***	3.604***	0.0943**
	(0.0419)	(0.180)	(0.0907)	(0.0405)
Observations	104	69	104	109
R-squared	0.027	0.120	0.001	0.002

Panel B: Effects on Well-being

VARIABLES	Happiness	Life Satisfaction	Stress	Altruism (Dictator Game)
Treatment	-0.568***	-0.434**	0.240	30.66**
	(0.176)	(0.168)	(0.167)	(13.68)
q-values	0.024	0.024	0.054	0.048
Constant	3.208***	3.434***	2.684***	54.90***
	(0.115)	(0.109)	(0.125)	(9.793)
Observations	103	103	102	104
R-squared	0.094	0.062	0.020	0.047

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The q-values (Benjamini et al., 2006) reported in each table treat all the outcomes in this table as multiple hypotheses being tested together. Standard errors are robust.

F Larger Scale Experiment

We conducted a demographic survey with a sample of $\sim$ 800 individuals in 2 NGO centers in Delhi in 2023 to alleviate the concern that the baseline patterns are specific to the sub-sample or because the survey was conducted during the pandemic. Table F.1 shows that the baseline patterns of low engagement both in terms of low dialogue and gaps in the number of connections in advice networks and overall networks continue to hold. The proportion of individuals who have never or rarely had a conversation about mental health and financial concerns is lower than before but still as high as 30%. We also separately measured the number of links for financial and mental health-related advice-taking and find that the average number of links in the former is even lower than that of advice-taking around mental health. We find that those who are more optimistic about their community's willingness to engage are more likely to have fewer gaps in connections between overall and financial networks. We also continue to find that individuals report a high willingness to engage and an even larger proportion underestimate how willing their community is to engage with them. This reassures us that the baseline patterns measured before the experiment continue to hold.

Next, we present the results from the implementation of a similar information treatment. Half of these individuals were chosen to be treated and given the same information as the treated individuals in the RCT with two crucial differences that make this a weaker replication of the original experiment. First, we provided individuals with information about the average willingness to engage in 2021. They were told that this is not information from 2023 and were asked to answer questions assuming this were true today.[34] Further, unlike the original

[34] This is because we were logistically constrained to contact individuals again after eliciting their willingness to engage and wished to run this additional experiment in the same round. As a result, we could only provide them with information collected previously.

treatment where individuals were told that the information we are providing is from their own community, in the additional sample, we told individuals that the information is collected from other similar communities of waste-pickers managed by the NGO in Delhi. This is because this center was not a part of our original sample and even if it were, the center composition would have changed due to COVID-induced migration.

Table F.2 shows the balance checks for this sample. We find that while there is balance in most variables, treatment and control groups are not similar in terms of the number of overall network connections, willingness to engage in mental health related dialogue, and the proportion whose belief about community's willingness to engage is lower than the delivered information. Since this is a larger sample, we can use post double selection Lasso (Belloni et al., 2014) to estimate the treatment effects. The effect on willingness to participate in savings groups and listening services is shown in Tables F.4 and F.6. We continue to find evidence that beliefs about community willingness to engage significantly affect network engagement. We also present robustness checks where we exclude those who were a part of the previous study or report having heard about their community's views on engagement via their peers in Tables F.5 and F.7. This is to ensure that the previous information delivery does not affect our findings. We find that the results are similar in direction, magnitude, and statistical significance whether these individuals are included or not.

The mechanisms behind this treatment are different from the original treatment in that individuals are provided information about other communities like theirs from two years ago. As a result, we find negative average effects on willingness to participate in savings groups. Disentangling this further, we find that those who are pessimistic about their community, are significantly less likely to engage when they are told that other communities are more willing to engage. Those who are optimistic about the community behave in the opposite manner. This pattern also holds for listening services even though the negative average effect is not significant. These effects can arise if those who believe that their community is not very willing to engage do not treat the new information as a positive signal about their own community but instead compare their community with other seemingly more supportive communities and draw a negative inference. This highlights the importance of the type of reference group whose beliefs are conveyed, as beliefs about peers can be updated either directly or through relative comparisons.

Regardless of how the information affects individual engagement, these results confirm that information about own community's willingness to engage (relative to others, in this case) can affect own willingness to engage with the network. Not only this, despite the weaker treatment, as Table F.8 shows, we detect a large increase in financial contributions made by the participants for setting up savings groups and listening services. These contributions are significantly higher than the control group for treated participants who have lower initial predictions about their community's willingness to engage with them i.e., those who think that not many people from their community would be willing to engage with them.[35] This suggests that while they may be less willing to engage upon hearing the information, they are more willing to finance avenues for such interactions to be set up in future. We are planning to use these collected funds to set up avenues for informal interactions (such as savings groups) with the help of the NGO.

[35] Note that this outcome is only accurately collected for the second center which is why the subsample here is close to half of the entire sample.

F.1 Results from the Larger Scale Experiment

Table F.1: Summary Statistics

	(1)	
Age	34.03	(12.35)
Female	0.549	(0.498)
Income	6780.9	(3142.1)
Degree (Overall)	4.655	(2.077)
Degree (MH Advice)	2.431	(1.310)
Degree (FC Advice)	2.216	(1.324)
Happiness (1-5)	3.326	(1.140)
Willingness to Talk (FC)	0.898	(0.303)
Willingness to Talk (MH)	0.880	(0.325)
Beliefs (MH; 0-10)	3.753	(1.800)
Beliefs (FC; 0-10)	3.571	(1.864)
Dialogue (MH; 1-4)	2.930	(1.011)
MH talk (Never)	0.113	(0.316)
MH talk (Rarely)	0.210	(0.407)
MH talk (Sometimes)	0.312	(0.464)
MH talk (Often)	0.365	(0.482)
Dialogue (FC; 1-4)	2.889	(1.010)
FC talk (Never)	0.119	(0.324)
FC talk (Rarely)	0.215	(0.411)
FC talk (Sometimes)	0.325	(0.469)
FC talk (Often)	0.341	(0.474)
Observations	791	

mean coefficients; sd in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The above table shows the summary statistics (mean and standard deviation) for various demographic characteristics of interest for the additional sample in 2023. "MH" refers to mental health concerns and "FC" refers to financial concerns.

Table F.2: Balance Tests for the Full Sample in 2023

	(1) Control	(2) Treatment	(3) (1) vs. (2), p-value
Female	0.558	0.539	0.591
Age	33.901	34.144	0.783
Income	6911.429	6657.143	0.256
Degree (FC Advice)	2.275	2.160	0.222
Degree (MH Advice)	2.410	2.451	0.666
Degree (Overall)	4.491	4.810	0.031
Dialogue (MH; 1-4)	2.919	2.941	0.767
Dialogue (FC; 1-4)	2.894	2.884	0.898
Willingness to Talk (MH)	0.857	0.901	0.056
Willingness to Talk (FC)	0.881	0.914	0.125
Beliefs (MH; 0-10)	3.678	3.825	0.251
Beliefs (FC; 0-10)	3.564	3.579	0.909
Underestimators (MH)	0.823	0.771	0.067
N	385	406	

Notes: This table presents the results of balance tests between the treatment and control group for the entire replication sample. The first two columns show the means of the variables for the Control and Treatment Group and the third column shows the p values for the two-sided test that the two means are equal. "MH" refers to mental health concerns and "FC" refers to financial concerns. Underestimators is a binary variable equal to 1 if the individual's belief about proportion of community willing to engage is less than six i.e. the information provided to the treatment group. *Degree* refers to the number of connections

Table F.3: Balance Tests for Center 2 in 2023

	(1) Control	(2) Treatment	(3) (1) vs. (2), p-value
Female	0.437	0.427	0.856
Age	30.934	29.653	0.266
Income	8166.887	8370.000	0.593
Degree (FC Advice)	2.318	2.153	0.308
Degree (MH Advice)	2.616	2.567	0.757
Degree (Overall)	4.954	4.767	0.478
Dialogue (MH; 1-4)	2.722	2.693	0.807
Dialogue (FC; 1-4)	2.755	2.660	0.419
Willingness to Talk (MH)	0.874	0.927	0.129
Willingness to Talk (FC)	0.914	0.927	0.684
Beliefs (MH; 0-10)	3.954	3.960	0.975
Beliefs (FC; 0-10)	3.411	3.547	0.527
Underestimator (MH)	0.755	0.733	0.668
N	156	156	

Notes: This table presents the results of balance tests between the treatment and control group for the second center in the replication sample. We present these balance tests separately as the mental health fund outcome was only accurately measured for this center. The first two columns show the means of the variables for the Control and Treatment Group and the third column shows the p values for the two-sided test that the two means are equal. "MH" refers to mental health concerns and "FC" refers to financial concerns. Underestimators is a binary variable equal to 1 if the individual's belief about proportion of community willing to engage is less than six i.e. the information provided to the treatment group. *Degree* refers to the number of connections

Table F.5: Replication Exercise: Effect on Willingness to Participate in Savings Groups (Robust)

VARIABLES	(1) Savings Group	(2) Savings Group	(3) Savings Group
Treatment	-0.0759**	-0.0748**	-0.0806***
	(0.0297)	(0.0297)	(0.0290)
Beliefs (MH)		-0.0400	
		(0.0255)	
Treatment x Beliefs (MH)		-0.00215	
		(0.0277)	
Beliefs (FC)			0.0709***
			(0.0195)
Treatment x Beliefs (FC)			0.0573**
			(0.0289)
Constant	0.944***	0.968***	0.950***
	(0.0327)	(0.0586)	(0.0443)
Observations	696	696	696
Number of groups	0	0	0

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table shows regression results from the replication sample in 2023, excluding those who were a part of the previous study or report having heard about their community's views on engagement via their peers. We use post double selection Lasso (Belloni et al., 2014) with robust standard errors to estimate the treatment effects where we select from a range of baseline variables such as gender, age, income, number of network connections, dialogue intensity, and willingness to engage. Column 1 shows the main treatment effects while Columns 2 and 3 interact the treatment indicator with beliefs. "Beliefs" refers to the respondent's prediction (standardized) about how many community members would be willing to engage around mental or financial concerns. "MH" refers to mental health concerns and "FC" refers to financial concerns.

Table F.4: Replication Exercise: Effect on Willingness to Participate in Savings Groups

VARIABLES	(1) Savings Group	(2) Savings Group	(3) Savings Group
Treatment	-0.0693**	-0.0726***	-0.0712***
	(0.0270)	(0.0272)	(0.0264)
Beliefs (MH)		-0.0311	
		(0.0221)	
Treatment x Beliefs (MH)		-0.0107	
		(0.0247)	
Beliefs (FC)			0.0616***
			(0.0175)
Treatment x Beliefs (FC)			0.0592**
			(0.0266)
Constant	0.882***	0.976***	0.891***
	(0.0334)	(0.0545)	(0.0451)
Observations	789	789	789
Number of groups	0	0	0

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table shows regression results from the replication sample in 2023. We use post double selection Lasso (Belloni et al., 2014) with robust standard errors to estimate the treatment effects where we select from a range of baseline variables such as gender, age, income, number of network connections, dialogue intensity, and willingness to engage. Column 1 shows the main treatment effects while Columns 2 and 3 interact the treatment indicator with beliefs. "Beliefs" refers to the respondent's prediction (standardized) about how many community members would be willing to engage around mental or financial concerns. "MH" refers to mental health concerns and "FC" refers to financial concerns.

Table F.6: Replication Exercise: Effect on Willingness to Participate in Listening Service

VARIABLES	(1) Listening Service	(2) Listening Service	(3) Listening Service
Treatment	-0.0442	-0.0444	-0.0430
	(0.0323)	(0.0324)	(0.0319)
Beliefs (MH)		-0.0677***	
		(0.0253)	
Treatment x Beliefs (MH)		0.0287	
		(0.0304)	
Beliefs (FC)			0.0213
			(0.0230)
Treatment x Beliefs (FC)			0.0953***
			(0.0324)
Constant	0.730***	0.789***	0.695***
	(0.0226)	(0.0609)	(0.0523)
Observations	789	789	789
Number of groups	0	0	0

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table shows regression results from the replication sample in 2023. We use post double selection Lasso (Belloni et al., 2014) with robust standard errors to estimate the treatment effects where we select from a range of baseline variables such as gender, age, income, number of network connections, dialogue intensity, and willingness to engage. Column 1 shows the main treatment effects while Columns 2 and 3 interact the treatment indicator with beliefs. "Beliefs" refers to the respondent's prediction (standardized) about how many community members would be willing to engage around mental or financial concerns. "MH" refers to mental health concerns and "FC" refers to financial concerns.

Table F.7: Replication Exercise: Effect on Willingness to Participate in Listening Service (Robust)

VARIABLES	(1) Listening Service	(2) Listening Service	(3) Listening Service
Treatment	-0.0290	-0.0270	-0.0316
	(0.0345)	(0.0345)	(0.0342)
Beliefs (MH)		-0.0747***	
		(0.0283)	
Treatment x Beliefs (MH)		0.0241	
		(0.0326)	
Beliefs (FC)			0.0250
			(0.0254)
Treatment x Beliefs (FC)			0.0957***
			(0.0348)
Constant	0.721***	0.818***	0.718***
	(0.0248)	(0.0673)	(0.0437)
Observations	696	696	696
Number of groups	0	0	0

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows regression results from the replication sample in 2023, excluding those who were a part of the previous study or report having heard about their community's views on engagement via their peers. We use post double selection Lasso (Belloni et al., 2014) with robust standard errors to estimate the treatment effects where we select from a range of baseline variables such as gender, age, income, number of network connections, dialogue intensity, and willingness to engage. Column 1 shows the main treatment effects while Columns 2 and 3 interact the treatment indicator with beliefs. "Beliefs" refers to the respondent's prediction (standardized) about how many community members would be willing to engage around mental or financial concerns. "MH" refers to mental health concerns and "FC" refers to financial concerns.

Table F.8: Replication Exercise: Effect on Contribution to Set up Savings Groups and Listening Services

VARIABLES	(1) Contribution (in Rupees.)	(2) Contribution (in Rupees.)	(3) Contribution (in Rupees.)	(4) Contribution (in Rupees.)
Treatment	2.468	2.916	2.334	2.632
	(2.299)	(2.415)	(2.272)	(2.330)
Beliefs (MH)		1.020		
		(1.281)		
Treatment x MH Beliefs		-4.470*		
		(2.520)		
Beliefs (FC)			0.414	
			(0.901)	
Treatment x FH Beliefs			-3.409**	
			(1.698)	
Beliefs (Average)				1.030
				(1.349)
Treatment x Avg. Beliefs				-5.158**
				(2.456)
Constant	16.96***	16.84***	17.00***	16.95***
	(1.364)	(1.359)	(1.362)	(1.366)
Observations	312	312	312	312
R-squared	0.004	0.019	0.014	0.023

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

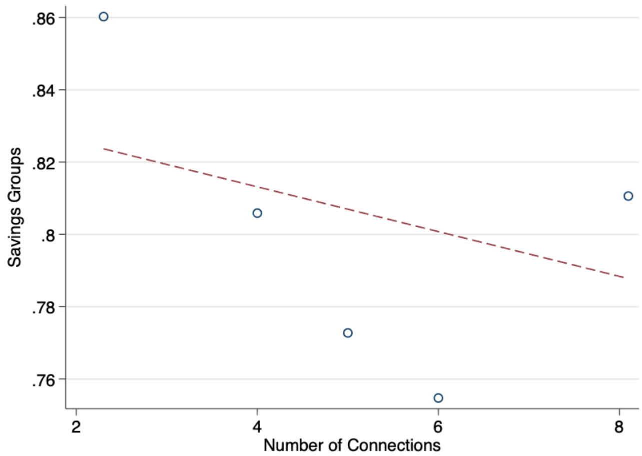
Notes: This table shows regression results from one center in the replication sample in 2023 for which this outcome was accurately measured, with robust standard errors. Column 1 shows the main treatment effects while Columns 2, 3, and 4 interact the treatment indicator with individual beliefs. "Beliefs" refers to the respondent's prediction (standardized) about how many community members would be willing to engage around mental or financial concerns. "MH" refers to mental health concerns and "FC" refers to financial concerns, and "Average" is the average of the two.

G Theory

G.1 Evidence for Model Assumptions

The following figures show correlations between degree centrality and survey measures of engagement with the network.

Figure G.1: Correlation between Degree and Willingness to Participate in Savings Groups



Notes: The figure shows the correlation between willingness to participate in savings group and the degree centrality of the agent using a binscatter with 5 quantiles.

G.2 Proof for Lemma 1

Consider the distribution function of the probability mass function $\widetilde{P}(d)$

$$\widetilde{\mathbf{P}}(\bar{\mathbf{d}}) = \sum_{d=1}^{d=\bar{d}} \widetilde{P}(d)$$

We first prove that the CDF corresponding to P is first order stochastically dominated by the CDF corresponding to $\widetilde{P}$. Consider $\bar{d} \leq E[d]$. In this case, $\sum_{d=1}^{d=\bar{d}} \widetilde{P}(d) = \sum_{d=1}^{d=\bar{d}} \frac{d}{E[d]} P(d) \leq \sum_{d=1}^{d=\bar{d}} P(d)$ since $\frac{d}{E(d)} \leq 1$.

When $\bar{d} > E[d]$, then $\sum_{d=1}^{d=\bar{d}} \widetilde{P}(d) = 1 - \sum_{\bar{d}}^{d=d_{max}} \widetilde{P}(d)$. The second term can be written as $\sum_{\bar{d}}^{d=d_{max}} \widetilde{P}(d) = \sum_{\bar{d}}^{d=d_{max}} \frac{d}{E(d)} P(d) > \sum_{\bar{d}}^{d=d_{max}} P(d)$ since $\frac{d}{E(d)} > 1$. So, $\sum_{d=1}^{d=\bar{d}} \widetilde{P}(d) < \sum_{d=1}^{d=\bar{d}} P(d)$ in this case as well.

Hence, $\widetilde{\mathbf{P}}(\bar{\mathbf{d}})$ first order stochastically dominates $\mathbf{P}(\bar{\mathbf{d}})$.

We can rewrite

$$\sum_{d} \widetilde{P}(d) H(c(d,a)) = H(c(d_{max}, a)) - \sum_{d=1}^{d=d_{max}-1} \widetilde{\mathbf{P}}(\mathbf{d})(H(c(d+1,a)) - H(c(d,a)))$$

and similarly rewrite

$$\sum_{d} P(d) H(c(d,a)) = H(c(d_{max}, a)) - \sum_{d=1}^{d=d_{max}-1} \mathbf{P}(\mathbf{d})(H(c(d+1,a)) - H(c(d,a)))$$

Given that H is weakly increasing in degree d and $\widetilde{\mathbf{P}}(\mathbf{d}) < \mathbf{P}(\mathbf{d})$, this implies that-

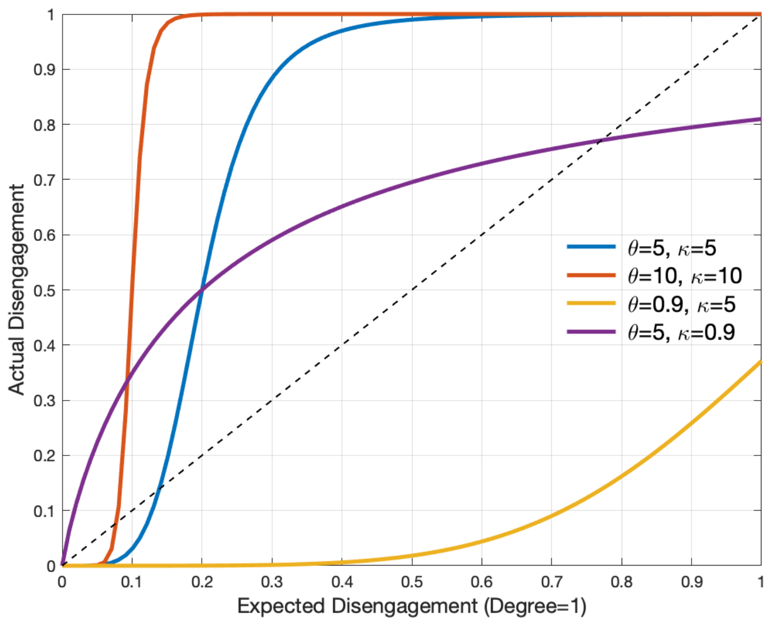
$$\sum_d \widetilde{P}(d)H(c(d,a)) \geq \sum_d P(d)H(c(d,a))$$

G.3 Proof for Proposition 1

If c and consequently H is strictly increasing in degree, Lemma 1 implies that $\widetilde{E}[H(c(d,z'))] > E[H(c(d,z')] \; \forall z' \in (0,1]$. This means that for any fixed point z that satisfies Equation 2, $\widetilde{E}[z] > z$. Given that we have assumed H to be such that there are three equilibria out of which 1 is unstable, this implies that the fixed points a_l and a_h solving equation 2 will not solve equation 3 and (b) the fixed points $\widetilde{a}_h$ and $\widetilde{a}_l$ that solve 3 will be such that $\widetilde{a}_l < a_l$ and $\widetilde{a}_h > a_h$. The intuition for Case 2 is analogous.
Click to go back.

H Structural Estimation

Figure H.1: Different shapes of $H(\theta, \kappa)$ for various values of θ and κ



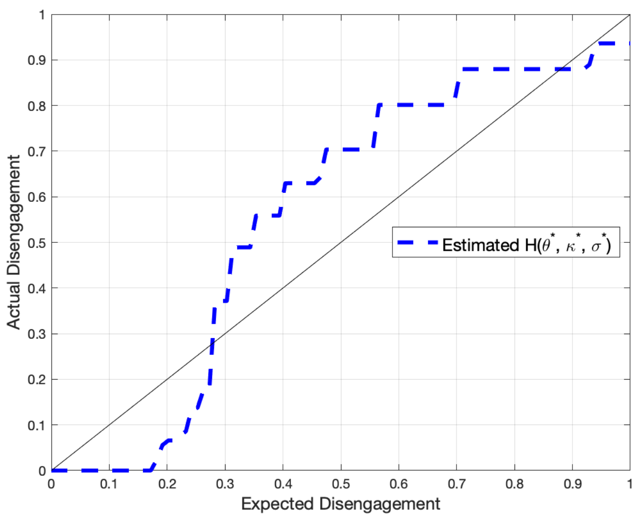
Notes: The figure plots the benefit distribution H distributed log-logistic with shape parameter κ and strategic complementarity parameter θ. Degree is normalised to be 1 in this example.

Table H.1: Structural Estimation Prediction Checks for various measures of Engagement

Variable	Objective Function	Predicted Mean (Treated)	Actual Mean (Treated)	Predicted Mean (Control)	Actual Mean (Control)
Listening Service	0.0005	0.309	0.299	0.440	0.460
Savings Groups	0.0012	0.217	0.187	0.318	0.333
Community Engagement	0.0001	0.201	0.200	0.332	0.338

Notes: This table shows the model fit for the estimated values of θ^* and κ^* for three different choice of outcomes. The objective function is computed at the estimated parameters and is equal to the sum of the squared percentage gap between the mean of the treatment and control groups predicted by the model and in the data respectively.

Figure H.2: Actual and Estimated Disengagement (assuming a Logistic Distribution)



Notes: The figure plots actual disengagement a and expected disengagement $\sum_d \widetilde{P}(d)(H(c(d,a)))$ at the estimated θ^*, κ^*, and σ^* in blue. Here, we assume a logistic distribution for H during the estimation so that $H(c(d,a)) = 1/(1 + e^{((-\theta * d * a) + \kappa))/\sigma})$.

Click to go back.