Peer Effects in Social Networks: Evidence from an Entrepreneurship Experiment [†]

Vatsal Khandelwal [§]

Juni Singh [¶]

Abstract

We implement a randomized entrepreneurship program to study how peer effects vary with network position and whether peers can be leveraged as bridges to useful social ties. We find that socially close and more connected peers generate short-term motivational gains, while close but less connected peers generate long-term benefits. The average long-term peer effect is close to zero, masking substantial heterogeneity: effects are largest for those matched with the least connected peers, declining by 0.06σ per connection. Sharing network contacts leads to no additional gains as contact sharing is concentrated within social groups, showing that bridging fails under high homophily.

JEL Codes: D85, C93, L26, L14, O12, Z13.

Keywords: Social networks; peer effects; social distance; entrepreneurship; weak ties.

[†]We thank the Weiss Fund for funding the fieldwork for this project. We also thank our partner organizations, JPAL and Governance Lab, as well as our enumerators for implementing the project. We are very grateful to Marina Agranov, Abhijit Banerjee, Michael Callen, Margherita Comola, Marcel Fafchamps, Horacio Larreguy, Sylvie Lambert, Karen Macours, Mahreen Mahmud, Rohini Pande, Simon Quinn, Simone Schaner, Charles Sprenger, Alexander Teytelboym, Verena Wiedemann, and Christopher Woodruff for their helpful feedback. The field experiment and endline survey are registered as AEARCTR-0010081 on the AEA RCT Registry and have received IRB approval both at Caltech and the University of Oxford.

[§]University of Exeter. v.khandelwal@exeter.ac.uk

[¶]Caltech, World Bank. jsingh2@worldbank.org

1 Introduction

Peer effects are documented in a wide range of contexts (Calvó-Armengol et al. 2009, Duflo et al. 2011, Banerjee et al. 2013, Cai et al. 2015, Field et al. 2016, Beaman et al. 2021), but the underlying mechanisms are less well understood (Bramoullé et al. 2020). Linear-in-means models used to estimate peer effects (Manski 1993) often place an equal weight on all peers even though peer characteristics can differentially affect outcomes in consequential ways. For example, popular peers may be effective role models in the short term, while socially close peers may be better collaborators in the long term. Further, peers not only have direct effects but can also help indirectly by connecting individuals to their other contacts. The relative importance of these "weak ties" (Granovetter 1973), and the extent to which they can be leveraged to generate additional effects, remains largely unexplored.[1]

In this paper, we provide causal evidence on how the proximity and popularity of the peer affect individual outcomes, how these effects evolve over time, and the extent to which peers can be leveraged as bridges to broader social networks. To do so, we implement a randomized entrepreneurship program in rural Nepal that induces peer interactions, experimentally varying both the identity of the assigned peer and whether peers share their network contacts. We find that while social proximity generates both short-term and long-term benefits, peer popularity can provide short-term motivational gains, particularly among socially close peers, but is detrimental in the longer term. We also show that bridging is difficult to engineer in settings with high homophily (i.e., preference for interacting with similar people) as contact sharing is concentrated among pre-existing social groups. By distinguishing the *relative* advantages of peers as motivators, collaborators, and connectors, we show that peer composition is a low-cost design choice that can substantially affect program effectiveness.

Causal evidence on how peer identity affects outcomes has been scarce, in part because development programs often pair peers endogenously, and in part because detailed network data are rarely available. To proceed, we first mapped the social networks of 2,840 women across 31 villages in Nepal and then randomized a subset to attend an intensive entrepreneurship training program.[2] While individuals in pure control villages did not attend the program, individuals in treated villages were allocated into one of three groups. They either completed the program alone, attended with another woman assigned by the experimenter, or did not attend at all. Among those attending in pairs, we additionally randomized the implementation of a 'connections' treatment in which trainees were encouraged to share their social contacts with each other and think of ways these can help them set up a business.

[1] The role of weak ties in connecting individuals to the broader social network also relates to the concept of "bridging" social capital introduced in Putnam (2000).

[2] Our main network measure captures all women in the village with whom an individual reports interacting in any capacity, including visiting homes, giving or taking advice, going to a temple together, or contacting during a health emergency.

We leverage the design to both provide reduced-form evidence and estimate peer effects. First, we exploit the variation in whether women attend the program alone versus with another woman in their village to assess whether social support can improve outcomes.[3] Second, we leverage whether individuals attended the connections treatment to compare the direct value of being trained with peers to the indirect value of accessing each other's social contacts. This allows us to assess whether increasing the perceived benefit of linking to new contacts and reducing the cost of reaching them can improve outcomes. Finally, we leverage the variation in the identity of the peer to study whether the treatment effect differs depending on their network position. In particular, we exploit variation in social distance (i.e., the number of network steps between two individuals) and centrality (i.e., the number of direct connections an individual has within the village network) across two survey waves to estimate heterogeneous peer effects. Our setting is well-suited for studying peer interactions, as social networks are sparse[4] and exhibit a strong community structure with high homophily, suggesting that networking opportunities may be limited.[5]

We first measure short-term outcomes immediately after the training program. Training alone significantly increases the willingness to seek additional resources post-training, by 0.24 standard deviations, but effects on knowledge, aspirations, and business intentions are not statistically significant. Being paired also does not lead to additional gains relative to training alone on average. However, the null average effect of the peer masks substantial heterogeneity. Pairing with a socially close and relatively more central peer significantly improves aspirations ($+0.74\sigma$) and willingness to take up additional resources ($+0.38\sigma$) relative to the control group – indicating the complementary short-term benefits of social proximity and network connectedness.[6] [7] Both effects are significantly larger than those of other pair types, and the take-up effect also significantly exceeds that of training alone. Consistent with the effects of role models, we also find that popular peers (i.e., in the highest quartile of centrality in their village) act as motivators. Those matched with popular peers have significantly higher aspirations compared to other pair types.

In the connections treatment, we find that individuals share 6.5 village-level contacts on average with their matched peer (about two-thirds of the combined network size of both peers). Those with greater social distance or a larger gap in connections share significantly more contacts. While this is promising, we find that individuals share significantly fewer contacts

[3] Peers can provide support through skill complementarities, motivation, and risk-sharing, especially since education, aspirations, and risk aversion are correlated with opening a business in our baseline data.

[4] These networks have an average density of 5.5%, i.e., only 5.5% of total possible network links are present. This contrasts with the average density of 11.9% in villages in Karnataka in Banerjee et al. (2024).

[5] Links with individuals of the same social group, i.e., caste, are also 21.5 percentage points higher, on average, than the proportion of same-caste members in the village population, indicating high levels of homophily.

[6] We define a person as socially close when their network distance is less than or equal to 2, i.e., strictly lower than the median social distance equal to 3.

[7] Heterogeneity along these dimensions was pre-registered in our analysis plan. Details on how we depart from our pre-analysis plan are provided in Section 3.

with those of a different social group.[8] Even among same-group peers, pairs with higher homophily (in-group linking bias) share fewer contacts. Consequently, sharing contacts via the connections treatment does not have a significant impact on short-term outcomes. This is despite network structures that make bridging valuable: 73% of pairs in this treatment arm do not share any friends in common. Taken together, these results indicate that even when structural conditions make weak ties valuable, homophily limits bridging across social boundaries.

We conducted a follow-up survey one year after the intervention. Those trained alone are significantly more likely to have taken steps to open a business, but those matched with a peer show no significant average effect. Decomposing this by peer identity as before, we find that those matched with a socially close peer have a significantly higher treatment effect on average outcomes compared to those matched with a socially distant peer. However, unlike the short-term results, those matched with more central peers are significantly *less* likely to have taken steps to open a business compared to those matched with less central peers. This suggests that peer centrality has different short-term and long-term effects, while social distance has similar effects over time.[9] We also detect positive spillover effects among control individuals within treated villages, with significantly higher effects for those with at least one network connection assigned to treatment.

We disentangle mechanisms and find that the short-term benefits of central peers are driven by motivation—encouragement is the most commonly reported benefit of pairing, and those matched with more central peers are significantly more likely to report feeling encouraged. Socially close pairs, by contrast, collaborate better during training (performing significantly better in a joint business plan exercise) and are more likely to report wanting to meet and start a business together. One year later, 31% of paired trainees report contacting their matched peer for advice, money, or business discussions, but those matched with more central peers are significantly less likely to discuss business with others in the village, and pairs across social boundaries are significantly less likely to have remained in contact. This suggests that any short-term motivational gains from central peers do not translate into sustained interactions, which instead occur mainly among peers of similar social status.

Next, we use the field data to estimate peer effects by exploiting random pairing to eliminate endogenous sorting into pairs and by using data from two survey waves to mitigate the "reflection problem" (Manski 1993). We find that the average peer effect is close to zero, but this masks significant heterogeneity: the effect of a 1 standard deviation increase in a peer's outcome on long-term outcomes ranges from 0.35 standard deviations for the least connected peers to a negative effect for the most connected, declining by 0.06 standard deviations for each additional peer connection. This heterogeneity persists even after controlling for

[8]In this setting, social group is defined using information on the individual's caste.

[9]We complement these findings with machine learning techniques that show that peer network characteristics are consistently ranked among the top three predictors of outcomes, compared to peer demographics.

individual and peer characteristics. These patterns are consistent with the mechanisms documented above: the motivational benefits of central peers are immediate, while the costs of sustained interaction with busier peers accumulate over time.

Importantly, while pairings based on social distance can be easy to implement, the feasibility of pairing based on centrality remains uncertain, since network data can be hard to collect and there may be a limit to how many centrality-based matches we can construct within a fixed network. We find that the majority of our sample guesses their peers' degree centrality correctly within one standard deviation, implying that additional data collection may not be necessary. Moreover, as a policy implementation exercise, we simulate 10,000 counterfactual reassignments of peer pairings within each village in our dataset and plot the distributions of the average centrality gap by village. The resulting distributions show wide dispersion, indicating that strategic pairings based on centrality are indeed feasible and can generate high returns.

Related Literature: Our paper contributes to the existing literature in three ways.

First, we contribute to the large literature on peer effects in economics by showing that their magnitude and sign vary systematically with peer identity and change over time. This heterogeneity helps identify the mechanisms through which peers influence outcomes—motivation in the short term, sustained interaction in the long term. We also show that peer assignment can be a low-cost lever that can substantially affect the effectiveness of development policies. While peer effects have been measured in various development contexts, including the adoption of new technology (e.g. Beaman et al. (2021)), financial products (e.g. Banerjee et al. (2013), Cai et al. (2015)), education (e.g. Calvó-Armengol et al. (2009), Duflo et al. (2011)), and entrepreneurship (e.g. Lerner & Malmendier (2013), Field et al. (2016)), the majority of literature treats peers as homogeneous. [10] [11] [12] By focusing on heterogeneity in peer effects, we also contribute to the literature on role models (Bernard et al. 2026, Riley 2024, Macours & Vakis 2014, Beaman et al. 2012), by showing that while socially proximate and more connected peers can be role models, peer connectedness can also be disadvantageous in the longer term.

Second, our experimental design allows us to distinguish between two contrasting mechanisms through which social capital can be beneficial—mechanisms that have not previously been disentangled experimentally. Peers can directly influence outcomes, or they can serve as an indirect gateway to a broader social network, potentially generating additional peer effects.

[10] Breza (2016), Breza et al. (2019) provide a review of the literature on social networks and development.

[11] The limitation of the linear-in-means assumption has also been noted in the context of peer effects in education (Sacerdote 2011) and more generally in Boucher et al. (2024), who show that taking an average can be misleading. Instead, we focus on the weights that individuals place on different types of peers.

[12] Exceptions include a growing literature that leverages variation in the peer's network position to show how it can heterogeneously affect different types of behaviour, ranging from contract enforcement (Chandrasekhar et al. 2018), to savings (Breza & Chandrasekhar 2019), and the development of social skills (Zárate 2023).

We disentangle these direct and indirect channels using a novel 'connections' treatment and highlight the extent to which peer effects can be engineered. By testing this, our design contributes to the large literature on social capital (e.g., Granovetter (1973), Glaeser et al. (2002), Durlauf & Fafchamps (2005)) and in particular, to the literature on "bonding" and "bridging" social capital introduced by Putnam (2000). We find that encouraging network formation by facilitating 'bridging' may not be feasible in contexts with high homophily.

Third, we contribute to the literature on entrepreneurship and the role of peers in improving business outcomes (Field et al. 2010, De Mel et al. 2014, Field et al. 2016, Cai & Szeidl 2018, Vasilaky & Leonard 2018, Fafchamps & Quinn 2018, Brooks et al. 2018, Carranza et al. 2018, Campos et al. 2019, Asiedu et al. 2023, Vega-Redondo et al. 2023). Unlike most existing business training RCTs, which focus on existing entrepreneurs or individuals who expressed interest in starting a business (see McKenzie et al. (2021) for a review), we mainly focus on non-business-owning women in our villages regardless of entrepreneurial intent. This allows us to examine whether training and peer effects operate for a broader, less selected population. We also show that peer effects in entrepreneurship mask significant network-based heterogeneity in that there are optimal peer matches that can be leveraged to improve outcomes. Related to our paper, Vega-Redondo et al. (2023) document the role of peer heterogeneity by showing the effects of diverse and non-diverse peer groups on entrepreneurship and how this varies based on interaction format (virtual or in-person). In contrast, we hold the interaction format fixed and vary peer identity within the same social network, allowing us to measure how network position affects entrepreneurship over time.

The paper is organized as follows. Section 2 describes the context and experiment design. We discuss the reduced form and peer effects estimation strategy in Section 3. We present the short-term and long-term results of the intervention in Section 4 and discuss the mechanisms in Section 5. This is followed by Section 6, where we estimate peer effects, and discuss the feasibility of strategic pairing. Section 7 concludes.

2 Data

We first conducted a detailed survey and network elicitation with 2,840 women across 31 villages in rural Nepal in September-October 2021. We collected data on demographic characteristics (age, marital status, caste, education), willingness to start a business, reasons for not doing so, risk aversion, and aspirations related to agricultural investments, non-agricultural investments, and income.[13] In addition, we administered a network survey to elicit information about social networks, as in Banerjee et al. (2013). These questions measure with whom individuals report interacting in any capacity, including visiting homes, giving or taking advice, going to a temple, or contacting during a health emergency. Indi-

[13]Data on aspirations were collected as per the procedure in Bernard & Seyoum Taffesse (2014).

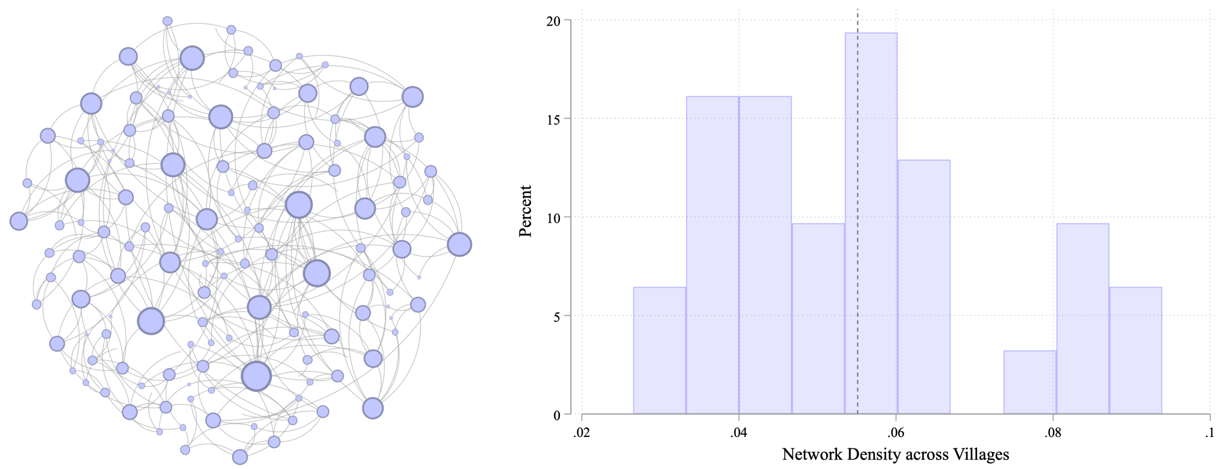


Figure 1: The figure on the left shows the social network in a village from our data. Larger nodes represent women with more connections. The figure on the right plots the network density for each village in our sample. Density is computed as the number of realized edges in the social network as a proportion of total possible edges. The number of total possible edges is given by $\frac{n(n-1)}{2}$ where n is the number of individuals in the network.

viduals were asked to list as many names as they liked. For the majority of the analysis, we treat links as undirected, i.e., individual i is assumed to be a friend of j if either mentions the other's name for any of the above interactions. We successfully elicited social networks for about 80% of village populations on average.[14][15]

2.1 Baseline

Baseline summary statistics are presented in Table A.1. The average age of women in our sample is 38, and 92% of them are married. Around 45% of our sample has no formal education. Roughly 22% of women report having opened businesses already, and 42% report a willingness to open a non-agricultural business. We find that 89% aspire to earn an income higher than their current income, while 23% aspire to spend more on non-agricultural business expenditures than their current investment. We also elicited risk preferences using a standard choice experiment involving a series of lotteries and a fixed payment. On average, women are risk-averse with a risk-aversion level of around 4.6, where 6 stands for very risk-averse and 1 stands for risk-loving.

[14] The village population is defined as the set of women who were either surveyed at baseline or listed as a network contact by a surveyed woman.

[15] Our main results are robust to correcting the centrality measure to account for our sampling strategy. These results are discussed in the supplementary appendix (Section I).

2.1.1 Network Statistics

Degree centrality is defined as the total number of undirected links an individual has in the village social network, i.e., the count of people who either named or were named by the individual as a friend. The average degree centrality in our sample is about 4.6.[16] The left panel of Figure 1 plots the social network of one of the villages in the sample, showing the heterogeneity of the number of connections reported. Table SA.1 shows that the number of connections is correlated with other demographic variables of interest, including education, caste, and marital status, with significantly fewer links among unmarried women, women with no education, or women of a lower caste.

Networks in these settings have two key features. First, networks are sparse, i.e., each person is connected to a small share of others in their village and most potential links are not present. We plot the density of networks, i.e., the number of observed links divided by the total possible links, computed for each village (Figure 1). The average network density across all villages is 5.5%, i.e., only 5.5% of links exist out of the total possible number of links given the number of individuals observed in each network. The sparsity persists even after we correct each individual's degree centrality (i.e., number of links) to account for our sampling strategy. Average density is 7% in that case. In comparison, the villages in rural Karnataka in Banerjee et al. (2024) have an average density of 11.9%.

Second, we find that these networks exhibit a community structure and a high level of homophily (i.e., in-group linking bias): intra-caste links are more common than links across castes, even after accounting for the number of same-caste members present in the village. On average, the proportion of same-caste links exceeds the proportion of same-caste village members by 21.5 percentage points. We measure caste homophily using the inbreeding homophily index (Currarini et al. 2009):

$$\text{Homophily} = \frac{\text{share of same-caste links} - \text{share of own caste in population}}{1 - \text{share of own caste in population}}.$$

Intuitively, the index measures what fraction of an individual's maximum possible in-group bias they actually exhibit, accounting for the fact that members of larger groups could appear biased even if they link randomly. The index ranges from -1 (maximum cross-type preference) through 0 (no excess same-type preference) to 1 (maximum in-group bias). The average inbreeding homophily in our sample is equal to 0.4.

Together, these findings suggest that the experiment induces social interactions that may not occur naturally.

[16] Degree centrality is measured at baseline for 95.3% of respondents. For the remaining 4.7%, who were not part of the baseline sample, degree is computed using their endline nominations, including any baseline nominations they received from others.

2.1.2 What are the barriers that prevent women from opening businesses?

We find that women who already own a business have a 50% higher median income than those with no business ownership. Correlating whether or not they have opened businesses with their baseline characteristics in Table SA.2, we find that those who own a business are younger, more educated, less risk-averse, and have higher aspirations for investment in non-agricultural business. This is also shown in Figure SA.1 that employs a Lasso regression (at various penalty levels) to show which variables correlate with whether or not a business has been opened already out of a list of demographic and network characteristics. We find that being educated, and especially having obtained a university education, matters the most in explaining the decision to have already opened a business.

When asked to give reasons for not having opened a business, 77% say that they do not have the skills or do not feel capable, while 24% say that they lack the financial ability. We correlate their willingness to open businesses with baseline demographics, number of network connections, aspirations, and other variables such as risk aversion. These results are shown in Table SA.3. We find that those who are older and more risk-averse are correlated with being less willing to open businesses. This is also reinforced by Figure SA.2 where we implement a Lasso regression, which selects age, risk-aversion, and education as the relevant correlates at high values of the penalty parameter. This suggests that risk-sharing and skill complementarities with peers could be helpful in setting up a business. Finally, those who have higher aspirations are more willing to open businesses, suggesting that peers can potentially be used to motivate and boost aspirations.

2.2 Experiment Design

We conducted the experiment in September 2022. The experiment consisted of a three-day entrepreneurship training program motivated by the SIYB module developed by the International Labour Organisation. The training typically lasted 3 hours per day, and individuals were given NPR 100/day (i.e., $0.7) for participation. The training focused on various topics, including defining entrepreneurship, listing the characteristics of a successful entrepreneur, learning how to build a business plan, learning about market scoping and market access, setting savings goals, and boosting aspirations using a video highlighting a successful female entrepreneur. Figure A.2 represents our two-step randomisation design.

First, we randomly allocated villages to Pure Control and Treatment. Following this, women in treatment villages were randomly allocated into one of the four groups as discussed below.

Control: Control individuals were not trained and remained untrained throughout the study. Their endline survey was conducted before any training took place in their village to minimize potential spillover contamination from interacting with trained participants. They serve as the spillover group in the one-year follow-up analysis.

Treatment 1: Trained Alone: Women in this group attended the training alone, i.e., the enumerator seated the participant alone without pairing them with another individual. Other women from the same village were present in the classroom, but there were no coordinated activities for these women throughout the training. These individuals completed all the interaction-based activities of the training alone, e.g., developing a business plan or playing an investment and savings game.

Treatment 2: Trained with a Peer: Women in this group attended the training paired with another person from their social network. Women were allocated a peer by the enumerator at the training centre, ensuring that this peer is not self-selected. These pairs completed all the interaction-based activities of the training with each other, e.g., developing a business plan or playing an investment and savings game.

Treatment 3: Trained with a Peer + Connections Module: As in Treatment 2, women in this group also attended the training with another person from their social network. However, in addition, the pair received a "connections module" that encouraged them to share their network contacts with each other and think about how network ties can be helpful in opening a business. To guide participants, we first highlighted three main ways in which networks can be helpful: i) Information-Sharing ii) Complementarity in skills, and iii) Risk Sharing, and then asked them to think about how they could help each other. Following this, the pair were asked to write down and share their network contacts that could potentially help them in opening up a business together, and think of ways each of these contacts could be helpful.

Conceptual Motivation for the Connections Module:

Conceptually, the connections module aims to test whether increasing the perceived benefits of linking and reducing the cost of reaching new contacts can induce useful interactions. Suppose individual i decides whether to interact with individual k with expected utility $E[U_{ik}] = \mu_{ik} - c_{ik}$, where μ_{ik} denotes the perceived benefit from interacting and c_{ik} is the cost of approaching k, including learning about them in the first place. In sparse, homophilous networks, μ_{ik} may be low and c_{ik} may be high. The module is designed to reveal who can be reached through the matched peer j and how these contacts might be useful, increasing μ_{ik} and lowering c_{ik} at the same time.

The scope for bridging depends on how much the pair's networks already overlap. When overlap is low, the peer provides access to genuinely new contacts; when overlap is high, the shared contacts are largely redundant. Given the sparsity of our networks (average degree ≈ 4.6 and average density of 5.5%), most experimentally assigned T3 pairs have non-overlapping neighborhoods: the average overlap of friends among T3 pairs is just 4%, with

73% of pairs sharing zero common contacts. This implies that the potential for bridging is high in our setting, but realizing this potential requires that the new contacts are useful and that individuals are willing to engage with them.

The connections module also allows us to compare the direct and indirect benefits of peers, comparing the direct role played by peer motivation and collaboration, for example, with the indirect role played by peers in facilitating additional useful interactions with others in the network.

2.3 Sampling and Implementation

Our design included three features to ensure internal validity. First, all individuals in treated villages were informed of their treatment status after arriving at the training centre. As a result, control and treated participants in our endline sample are not selected on the basis of their intentions to come to the training venue. Second, among those attending training with a peer, peer assignment occurred after individuals arrived at the training centre. This means that initial decline of invitations to participate prior to arrival at the centre cannot be correlated with peer type, since non-attendees were never assigned a peer. As a result, our main comparisons of interest, between those trained alone versus with peers and across peer types within the paired arms, are unaffected by selection concerns related to these dropouts. Third, for the control group, we aimed to conduct the endline survey before the training began for anyone in their village to minimize the risk of spillovers. The endline survey took place immediately on the third day of training for those who attended the training.

Selection: Out of 2,840 women surveyed at baseline, we excluded those already operating businesses and four villages where training could not be implemented ex-ante. We then randomized the remaining 1,970 women across 27 villages into four treatment arms. Nearly half of the randomized sample participated in the experiment, and 44% of those invited in treated villages attended the training venue. These attendance rates are consistent with those documented in the business training literature, where average take-up is approximately 65 percent even when training is offered for free, and rates as low as 39–51 percent are common even among samples screened for initial interest (McKenzie & Woodruff 2014). As expected, participation rates do not differ between those assigned to treatment and those in the within-village control group (Table B.3). This is because neither group was informed about their treatment status prior to arrival at the training centre. Dropout rates are also similar for those trained alone and those trained with a peer, which is our main comparison of interest. To evaluate representativeness of the sample for external validity, we assess baseline characteristics of dropouts and non-dropouts in treated villages. These findings are discussed in Section B.2 in the appendix.

We conducted a phone-based follow-up survey one year later with a subset of 580 women from the endline sample. Table B.5 shows that attrition between endline and follow-up is also not correlated with assignment to training. Non-response was primarily due to changes in phone numbers. Resource constraints prevented an in-person follow-up.

Compliance: We define compliance as a binary indicator equal to one if an individual's realized treatment status matched their assignment. Overall compliance in the endline sample is 68%, or 58% when excluding pure control villages, where compliance is mechanical. Non-compliance is two-sided: among those assigned to training who arrived at the venue, 82% were actually trained, while some within-village controls were included in the training to fill available spots. We do not find meaningful differences in baseline characteristics between compliers and non-compliers in treated villages (Table B.2). Given non-compliance, the main results report intention-to-treat (ITT) estimates. We also report LATE specifications that instrument actual treatment with treatment assignment; first-stage F-statistics are large in all specifications. Further details are provided in Appendix Section B.1.

Balance: We check for balance on several baseline variables within treated villages, including demographic, network, and business characteristics such as income, sources of income, age, education, caste, network connections (i.e., degree centrality), and aspirations, for all individuals in the endline and follow-up samples. Tables B.6 and B.7 show these results for the endline and follow-up samples, respectively.[17] Comparing within-village controls to each treatment arm (57 pairwise comparisons across 19 variables), 7 are significant at the 5% level, with most driven by degree centrality. Comparing across treatment arms only, only 3 are significant at the 5% level. We find similar results for the follow-up sample. This suggests that while degree centrality differs between controls and treated individuals, the treatment arms themselves are well-balanced.

To ensure that imbalance in degree centrality is not a concern, we check for differences in the degree distribution of control and treated individuals using Kolmogorov-Smirnov tests (Figure SA.3)[18]. We find that there are no significant differences in the distribution of network connections of individuals across the various arms. Further, we also later show that our key results are robust to any observed cases of imbalance by employing post-double selection Lasso wherever necessary (Belloni et al. 2014).

2.3.1 Peer Assignment and Randomization Inference

Peers were assigned on the spot by the research implementation team, with the goal of avoiding self-selection by 'randomly' pairing women with another attendee from the same village. We first assess whether the implemented assignments are as-good-as-random by conducting randomization inference on the entire sample of individuals who were actually

[17] Table SA.8 in the supplementary appendix shows endline balance checks including pure control villages and finds similar results.

[18] Tables and figures numbered with the prefix "SA" appear in the Supplementary Appendix, available online.

paired for the training. Specifically, we simulate multiple alternative peer assignments within each village and compute, for each assignment, a statistic capturing the total difference in characteristics between assigned pairs. We find that the observed value of the statistic from the actual pairing lies well within the distribution generated from these simulations, suggesting that the implemented peer assignments are statistically indistinguishable from random. These results are discussed in Section B.5.

Balance by Peer Type: We additionally check whether baseline characteristics of individuals differ by the type of peer assigned within the paired treatment arms. Tables SA.4–SA.7 compare individuals assigned to close versus distant peers and more versus less central peers in the endline and follow-up samples. Most characteristics are well-balanced across different pair types: 14 of 18 individual characteristics show no significant differences by social closeness with peer, and 15 of 18 by peer centrality in the endline sample. Importantly, income is balanced by peer centrality, as are willingness to start a business, caste, and most education categories. Some imbalances arise with respect to social closeness (e.g., individuals assigned to close peers have higher baseline income, p=0.010, and are younger, p=0.092) and by peer centrality (those assigned to more central peers have lower rates of higher education, p=0.002). Such correlations are not surprising: because network position reflects underlying demographic characteristics, baseline covariates are expected to correlate with the peer's network attributes even under random assignment. However, we present a variety of checks throughout the paper which highlight the impact of network attributes over and above the impact of any correlated baseline characteristics (see Section SA.1).

3 Estimation Strategy

3.1 Endline Outcomes

We measure the effect of the experiment on five endline outcomes collected after the training. These include the Knowledge Index, Business Aspirations Index, Business Index, Take-up Index, and Additional Steps, consistent with our pre-analysis plan.

The indices are constructed as follows: (1) **Knowledge Index**: We ask five short questions that measure knowledge gained during the course of the training, and take the proportion of correct responses to construct the knowledge index. We collect this measure for both treatment and control groups.[19] (2) **Business Aspirations Index**: We compute a measure of business-related aspirations comprising yearly non-agricultural investment, monthly income, and savings, elicited following the procedure in Bernard & Seyoum Taffesse (2014). We ask individuals the minimum and maximum of the relevant variables in their neighbourhood, the

[19] The questions are: (a) What do you understand by a business? (b) What characteristics are required to be a successful entrepreneur? (c) What do you mean by fixed assets? (d) What sector does a beauty parlour come under? (e) Above what break-even percentage does the business become risky?

current value, and what they aspire to in the corresponding time frame, and construct the weighted average of the individual's aspirations across these dimensions to create the index. (3) **Business Index**: This is constructed as a weighted average of responses to whether the individual is ready to invest in a business, whether they would submit a plan for a business competition, and how likely they are to start a business, capturing forward-looking intention and self-assessed likelihood of taking entrepreneurial action. (4) **Take-up Index**: This is a weighted average of hypothetical willingness to seek additional support. Specifically, we ask if individuals would attend additional paid training sessions or mentoring workshops in the next year, how much they would be willing to pay for each, and whether they are open to receiving advice from community members. (5) **Additional Steps**: This index captures concrete steps toward setting up a business, including whether the individual intends to open a savings account or take a loan for the business. In each case, we construct a weighted average (Anderson 2008) and normalize indices using the within village control group. [20]

With the exception of the Knowledge Index, which tests factual knowledge, these are self-reported measures capturing aspirations, intentions, and willingness rather than actual behaviors. This allows us to assess immediate changes in mindset and motivation following the training. Because each index aggregates several underlying variables, movements in these indices capture broad changes in intentions and mindset and are substantively meaningful.

One concern with these measures is that stated intentions may not translate into changes in actual behavior. However, we show that these endline measures are significantly correlated with follow-up behaviors one year later, as defined below (Table SA.9).

3.2 Follow-up Outcomes

Treatment effects on follow-up outcomes are measured one year later with a subset of the endline sample using phone surveys. These outcomes include whether or not the individual has opened a business, their monthly income, agricultural investments, agricultural profits, whether they have opened a new savings account, the amount of money they save, and whether they have taken a loan. In addition to these main economic outcomes, we also measure other outcomes, including income aspirations, whether individuals sign up for a potential commitment savings account from which funds cannot be withdrawn except for business purposes, record-keeping for agriculture, and other outcomes regarding community interactions around advice-taking and collaborations.

We construct four indices to capture follow-up outcomes.[21] First, the **Outcomes Index** includes monthly income and agricultural profits. Second, the **Steps Index** captures con-

[20] Variables entering the aspirations index and payment variables in the take-up index were winsorised at the 99th percentile.

[21] We rely on an intuitive categorization since the follow-up survey was not pre-registered.

crete actions taken toward setting up a business, including whether the individual reported opening a business, opening a new savings account, or joining a cooperative. We also include in this the level of business investments, agricultural investments, savings, and loans. Third, the **Mindset Index** captures forward-looking motivation and includes income aspirations, willingness to start a business, perceived self-efficacy in running a business, and willingness to open a (hypothetical) commitment savings account (both participation decision and amount committed), where the amount can only be withdrawn for business-related expenses. Fourth, the **Business Practices Index** includes whether individuals keep records for agriculture or for business, capturing improvements in organization. As before, all indices are normalized relative to the pure control group.[22]

3.3 Empirical Specification

3.3.1 Intent-to-treat Effects

We measure the immediate impact of the treatment using the main specification described below, where we compare treated individuals with control individuals within treated villages:

$$\begin{aligned} Y_i = \alpha &+ \beta_1 \, \text{Trained Alone}_i + \beta_2 \, \text{Trained With Peer}_i \\ &+ \beta_3 \, (\text{Trained With Peer+Connections Module})_i + \epsilon_{iv} \end{aligned}$$

Y_i is an outcome measure for individual i, *Trained Alone* is a binary variable that takes the value 1 if the individual was intended to be treated alone and 0 otherwise. Similarly, *Trained With Peer* is a binary variable that takes value 1 if the individual was intended to be treated with a peer and *Trained With Peer+Connections Module* is a binary variable that takes value 1 if the individual was intended to be treated with a peer and attend the additional connections module. The omitted category is the within-village control group, i.e., individuals in treated villages who were not assigned to training.

For the follow-up analysis, we include pure control villages and additionally control for an indicator variable for the spillover group (within-village controls in treated villages). Standard errors are clustered at the village level, when pure control villages are included and robust otherwise.

Local Average Treatment Effects: As a robustness exercise, we also estimate specifications that instrument actual treatment status with assignment to obtain local average treatment effects. These results are reported in the Supplementary Appendix (Section H).

[22] All monetary variables were winsorized at the 1st and 99th percentiles to exclude outliers. This includes monthly income, total savings, agricultural profit, agricultural investment, and income aspirations. In contrast, business investment and loan amount were winsorized at the 99th percentile only, since the majority of respondents reported zero values for these variables.

3.3.2 Impact of Training with different types of peers

We consider differences in outcomes for different pairs in treatments 2 and 3 compared to treatment 1. Our intention to leverage the random variation in the network identity of the peer and use it to study heterogeneity by social distance and network centrality was pre-registered in our pre-analysis plan.

We implement specifications using both continuous and discrete measurements of network statistics, as well as absolute and relative measures of peer centrality (i.e., comparisons with the matched peer's degree). Let d_{ij} be the network distance between i and j, i.e., the number of links needed by i to reach j, and let ϕ_i be the network centrality (e.g., number of connections) of agent i. First, we combine the paired treatment arms and interact the indicator for each with continuous measures of peer social distance and peer centrality, and implement the following specification.

$$
\begin{aligned}
Y_i = \alpha &+ \beta_1 \,\text{Trained Alone}_i + \beta_2 \,\text{Trained With Peer}_i \\
&+ \beta_3 \,(\text{Trained With Peer}_i \times d_{ij}) + \beta_4 \,(\text{Trained With Peer}_i \times \phi_j) + \epsilon_i
\end{aligned}
$$

Following this, we split individuals into whether or not they are socially close. We define a pair ij as socially close if the network distance between i and j is strictly less than the median social distance in the sample (median $d_{ij} = 3$), i.e., $d_{ij} \leq 2$. Similarly, we classify peers into two additional categories: if they have the same or fewer network connections ("Less Central") or more network connections ("More Central") than their matched peer. We then implement the following specification, where we regress outcomes on the indicator for being trained alone, trained with a peer, trained with a close peer, and trained with a more central peer.

$$
\begin{aligned}
Y_i = \alpha &+ \beta_1 \,\text{Trained Alone}_i + \beta_2 \,\text{Trained with Peer}_i \\
&+ \beta_3 \,\text{Trained with Close Peer}_i + \beta_4 \,\text{Trained with a More Central Peer}_i + \epsilon_i
\end{aligned}
$$

Finally, we classify each peer type into one of four categories: Close × Central, Close × Noncentral, Far × Central, and Far × Noncentral. For any pair ij, i is assigned to the category close-central if their matched close peer j has strictly higher degree centrality compared to i, i.e., $\phi_i - \phi_j < 0$ and i and j have social distance $d_{ij} \leq 2$. The remaining categories are collapsed into a single indicator, Other Pair Types.

$$
\begin{aligned}
Y_i = \alpha + \beta_1 \,\text{Trained Alone}_i + \beta_2 \,(\text{Close} \times \text{Central})_i \\
+ \beta_3 \,\text{Other Pair Types}_i + \beta_4 \,\phi_i + \epsilon_i
\end{aligned}
$$

Additional Controls: Because an individual who is more central is mechanically more likely to be paired with a relatively less central peer under random assignment, own centrality will tend to correlate with relative peer type. Accordingly, all regressions that use a relative measure of peer centrality (e.g., whether the peer is "More Central" than the individual) include controls for the individual's own centrality. Finally, in all regressions that compare peer types or analyze heterogeneity within the paired arms, we control for individuals who were intended to be paired but were not actually matched due to non-compliance.[23]

Robustness: We also show robustness of our main results to alternative measures of centrality, incomplete network sampling, and additional controls. This includes using eigenvector centrality (importance from being connected to other well-connected nodes) and indegree centrality (importance from receiving many links) as alternative measures, using an alternative measure of degree centrality that accounts for our sampling strategy and any unobserved links, and including controls for additional observable peer characteristics.

3.3.3 Peer Effects

In addition to the reduced form analysis, we identify peer effects by exploiting random pairing and the panel structure of our data, estimating a linear-in-means specification (Manski 1993) with heterogeneity based on network position. This is discussed in more detail in Section 6.

4 Results

Table 1 shows the impact of the intervention on endline outcomes. Compared to within-village controls, training alone significantly increases the take-up index by 0.24 standard deviations ($p<0.05$); this result remains significant after adjusting for multiple hypothesis testing ($q<0.10$). Training alone also leads to significantly more additional steps than training with a peer ($p<0.05$). Treatment effects on knowledge and aspirations are positive but not statistically significant. The paired treatments (with or without the connections module) do not significantly affect any outcome relative to within-village controls.[24]

IV specifications capturing local average treatment effects among compliers (Table SA.26) yield larger coefficients compared to ITT, with significant effects on the take-up index for both training alone ($+0.98\sigma$, $p<0.05$) and training with a peer ($+0.74\sigma$, $p<0.10$). Other outcomes show positive but statistically insignificant effects. As with the ITT estimates, pairing continues to have no additional impact beyond training alone, on average.

Using pure control villages as the comparison group (Table SA.19), all treatment arms show

[23] Peer-specific regressors (e.g., social distance, peer centrality) are set to zero for all individuals not originally assigned to Treatment 2 or Treatment 3 in specifications that include all treatment arms.

[24] Table SA.10 shows the types of businesses individuals report that they want to open after being trained, conditional on being ready to invest. We discuss these results further in Section 5 when we explore mechanisms.

significant improvements in knowledge (p<0.01), while training alone additionally improves business intentions (p<0.05) and take-up (p<0.10).The pattern that pairing provides no additional benefit over training alone holds in all specifications.

Table 1: Impact of the training on immediate outcomes

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Intention	(4) Steps Index	(5) Take-up Index
Trained alone	0.114	0.125	0.108	0.126	0.243**
	(0.0945)	(0.112)	(0.102)	(0.0986)	(0.0969)
Treatment with Peer	0.0209	0.108	-0.0314	-0.0753	0.0928
	(0.0957)	(0.115)	(0.105)	(0.106)	(0.105)
Treatment with Peer + Connections Module	0.122	0.0483	0.0735	0.0760	0.149
	(0.0963)	(0.122)	(0.111)	(0.103)	(0.105)
Constant	0	0	0	0	0
	(0.0754)	(0.0754)	(0.0754)	(0.0758)	(0.0756)
Observations	710	710	707	688	693
R-squared	0.004	0.002	0.003	0.007	0.009
p: Alone vs Peer	0.259	0.889	0.167	0.0378	0.113
p: Alone vs Peer+CM	0.921	0.545	0.743	0.595	0.321
p: Peer vs Peer+CM	0.231	0.642	0.338	0.137	0.584
q: Alone	0.301	0.301	0.301	0.301	0.0660
q: Peer	1	1	1	1	1
q: Peer+CM	1	1	1	1	1

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

Notes: This regression treats within-village controls as the base category. Standard errors are robust. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across the five outcome indices. Treatment effects on individual index components are discussed in the Supplementary Appendix (Section L).

4.1 Effect of the Connections Module

Next, we find that the connections module leads to no statistically significant gains relative to other arms. This is despite women assigned to this treatment sharing 6.5 contacts on average with each other and discussing how these connections can help them set up a business. Within this treatment arm, peers who shared more connections were also more likely to report wanting to start a business together, even after controlling for social closeness and peer centrality (Table E.6).

However, Table E.5 shows that while more connections were pooled among those with a larger social distance, higher gap between network degrees, and lower measured overlap in their social networks, significantly more connections were shared between same-caste pairs (about 1.5 additional contacts). In magnitude, this effect dominates other predictors of contact sharing. As a comparison, for instance, a social distance of five implies less than one additional pooled contact on average. A Lasso regression (Figure SA.5) confirms that caste similarity is the strongest predictor of the number of contacts pooled, followed by own degree centrality and education.

Consistent with these findings, we also find that individuals in castes with high homophily (i.e., social groups within the village that have an ingroup bias in terms of network links) share significantly fewer contacts (Table E.7).[25] Even among individuals in the same caste, those belonging to high-homophily castes share even fewer contacts with each other (p=0.01). These results are discussed in more detail in Section 5.

Combined, these patterns suggest that it may be difficult to encourage network interactions in settings with high homophily due to a reluctance to share contacts across caste boundaries.

4.2 Effects of Peer Identity

Now, we exploit variation in the identity of the matched peer in order to understand the conditions under which pairing can be helpful. We measure treatment effects using various measures of social distance and network centrality.

First, we interact continuous measures of social distance and peer centrality with an indicator for being in the paired treatment arms. We find in Table 2(a) that higher social distance among peers significantly reduces endline aspirations. A one unit increase in social distance reduces aspirations by 0.14 standard deviations (significant at 1%); this finding holds after correcting for multiple hypothesis testing (q<0.05) and is robust to post-double selection Lasso (Belloni et al. 2014), which flexibly controls for baseline covariates including income, age, education, caste, and network connections that could be correlated with an individual's closeness to others in the network (p<0.10; Table SA.23).

At the same time, we find evidence that social distance increases take-up (+0.05σ, p<0.10), though this effect does not survive correction for multiple hypothesis testing (q=0.24). Examining individual components (Table SA.45), we find that this effect is entirely driven by a higher willingness to seek advice from other women in the village about setting up a business. Being matched with close peers can reduce the need to seek advice externally, while contact with distant peers can prompt individuals to look beyond their paired partner for additional guidance. Being matched with more central peers, on the other hand, also increases the take-up index—for every one extra connection of the matched peer, the take-up index rises by 0.05 standard deviations; this effect is robust to baseline controls (PDS-Lasso p<0.01; Table SA.23). Unlike the effect of distance, the centrality effect holds across multiple subcomponents, including willingness to seek advice from other women, interest in mentoring from women entrepreneurs, and willingness to participate in a business training program (Table SA.45).

We then implement a similar specification where we use a discrete classification of social closeness as defined earlier and a relative comparison of an individual's own degree with

[25] Homophily is measured using the inbreeding homophily index (Currarini et al. 2009), as defined in Section 2.

their peer's. Consistent with the previous result, Table 2(b) shows that being trained with a close peer increases aspirations by 0.50 standard deviations (significant at 1%) relative to being trained with a more distant peerand the result holds after correcting for multiple hypothesis testing (q<0.05).[26] Moreover, we find that the total effect of being matched with a close peer on aspirations is significantly different from zero implying a positive treatment effect for this subgroup. The effect is larger than for those trained alone but the difference is not statistically significant (p=0.16). As suggested by the previous specification, the total effect of being paired with a more central peer on take-up is significantly different from zero (p=0.004).

Complementarity between social distance and centrality: While the above specifications highlight the importance of social distance and centrality, they do not account for complementarity between the two, i.e., whether it is better if the socially close peer is well connected. To test this, we implement a two-by-two categorization by social distance and centrality: socially close peers that are more central than the individual, socially close peers that are less central, distant peers that are more central, and distant peers that are less central. As shown in Table 2(c), those who are paired with a more central and close peer have significantly higher aspirations (+0.74σ relative to within-village controls), exceeding the other pairs (p=0.02) and those trained alone (p=0.06). This effect persists after accounting for multiple hypothesis testing (q<0.10). Other pair types do not differ significantly from the control group on aspirations.

This suggests that central peers motivate more effectively when they are also socially close. Consistent with this, as we will shortly discuss in Section 5, those who were matched with central peers were also significantly more likely to list encouragement as a reason why pairing was useful for them. It is important to note that “more central" here refers to relative centrality—whether the peer has higher degree than the individual—not absolute popularity.

Table 2(c) further shows that those matched with a more central and close peer have significantly higher take-up (+0.38σ relative to within-village controls), significantly higher compared to the other pairs (p=0.04); notably, this result is also robust to multiple testing corrections (q<0.01). Among socially close peers, those with more central peers also have significantly higher take-up than those with less central peers (p=0.04; Table SA.11).

Importantly, the key patterns of heterogeneity that we document, i.e., the negative effect of social distance on aspirations and the positive effects of close × central peers on aspirations and take-up are also robust to degree-corrected centrality measures that account for incomplete network sampling (Table SA.28). Finally, under a more conservative PDS-Lasso specification (Table SA.23), social distance continues to reduce aspirations and peer centrality continues to increase take-up, both remaining statistically significant. Notably,

[26] Note that in this additive specification, the “Close" coefficient represents the *additional* effect of closeness, not the total effect relative to control.

Table 2: Effects on endline outcomes by peer type

(a) By Social Distance and Peer Centrality

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Intention	(4) Steps Index	(5) Take-up Index
Trained alone	0.114	0.125	0.108	0.126	0.243**
	(0.0946)	(0.113)	(0.102)	(0.0987)	(0.0970)
Trained with Peer	0.420**	0.695**	0.176	0.0383	-0.158
	(0.171)	(0.274)	(0.219)	(0.188)	(0.215)
Trained with Peer × Peer distance	-0.0282	-0.135***	-0.0210	0.00379	0.0471*
	(0.0254)	(0.0438)	(0.0344)	(0.0292)	(0.0283)
Trained with Peer × Peer degree	-0.0190	-0.0295	0.00125	0.0105	0.0480**
	(0.0195)	(0.0270)	(0.0250)	(0.0199)	(0.0213)
Constant	0	0	0	0	0
	(0.0755)	(0.0755)	(0.0755)	(0.0759)	(0.0757)
Observations	695	695	692	673	679
R-squared	0.047	0.016	0.013	0.020	0.037
q: Alone	0.302	0.302	0.302	0.302	0.0670
q: Paired	0.0380	0.0380	0.532	1	0.532
q: Paired x Dist	0.366	0.0110	0.687	0.807	0.238
q: Paired x Deg	0.786	0.786	1	1	0.143

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

(b) By Social Closeness and More/Less Central

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Intention	(4) Steps Index	(5) Take-up Index
Trained alone	0.0857	0.112	0.0797	0.112	0.204**
	(0.0936)	(0.114)	(0.101)	(0.0988)	(0.0943)
Trained with Peer	0.198**	-0.125	0.0632	0.0818	0.171
	(0.0948)	(0.129)	(0.119)	(0.114)	(0.113)
Trained with Peer (Close)	-0.0400	0.501***	0.0145	-0.0337	-0.0542
	(0.0983)	(0.186)	(0.128)	(0.119)	(0.108)
Trained with Peer (More Central)	0.0240	0.108	0.0263	0.0319	0.174
	(0.0955)	(0.165)	(0.130)	(0.123)	(0.111)
Degree	0.0466***	0.00162	0.0523***	0.0262	0.0371***
	(0.0129)	(0.0199)	(0.0163)	(0.0160)	(0.0143)
Constant	-0.223**	-0.00778	-0.250**	-0.126	-0.178
	(0.102)	(0.114)	(0.114)	(0.112)	(0.112)
Observations	700	700	697	678	684
R-squared	0.061	0.019	0.027	0.023	0.037
p: Total Close = 0	0.183	0.0468	0.582	0.725	0.385
p: Total Central = 0	0.0481	0.903	0.503	0.371	0.00401
p: Total Close = T1	0.503	0.158	0.988	0.620	0.487
p: Total Central = T1	0.183	0.372	0.941	0.987	0.197
q: Alone	0.527	0.527	0.527	0.527	0.180
q: Paired	0.231	0.505	0.556	0.549	0.349
q: Close	1	0.0380	1	1	1
q: Central	1	1	1	1	1

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

(c) By Social Closeness Interacted with More/Less Central

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Intention	(4) Steps Index	(5) Take-up Index
Trained alone	0.0861	0.107	0.0793	0.112	0.207**
	(0.0936)	(0.114)	(0.101)	(0.0987)	(0.0942)
Close x More Central	0.281*	0.744**	0.236	0.0899	0.375***
	(0.146)	(0.332)	(0.154)	(0.165)	(0.0941)
Other Pair Types	0.177**	-0.0286	0.0495	0.0816	0.195**
	(0.0874)	(0.104)	(0.100)	(0.0977)	(0.0975)
Degree	0.0459***	0.0123	0.0530***	0.0246*	0.0313**
	(0.0122)	(0.0196)	(0.0153)	(0.0146)	(0.0135)
Constant	-0.220**	-0.0590	-0.254**	-0.118	-0.150
	(0.0999)	(0.114)	(0.110)	(0.107)	(0.108)
Observations	700	700	697	678	684
R-squared	0.061	0.023	0.028	0.023	0.036
p: Close+Central vs Other	0.440	0.0195	0.219	0.959	0.0357
p: Close+Central vs T1	0.158	0.0566	0.303	0.887	0.0387
q: Close+Central	0.0680	0.0540	0.101	0.189	0.00100
q: Other Pairs	0.131	0.889	0.873	0.677	0.131

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

Notes: These regressions treat within-village controls as the base category and include an indicator for those who were intended to be paired but remained unmatched. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality. Standard errors are robust. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes.

the continuous effect of peer centrality on aspirations is negative (-0.04σ, $p<0.10$), which reinforces the complementarity documented above. Centrality averaged across close and distant peers does not improve aspirations, because the motivational benefit requires social proximity.

4.2.1 Peer Characteristics and Homophily

We find that training with a central and close peer outperforms other pair types in terms of the effect on aspirations and take-up, even after controlling for other peer characteristics and for the effects of homophily, i.e., the effect of being paired with peers who are similar in baseline characteristics. Moreover, once we account for homophily, the effect of being matched with a more central and close peer also outperforms being trained alone for take-up and business outcomes ($p<0.10$). These results are shown in Tables E.1 and E.2.

To show this, we first construct a similarity index to assess if the similarity between the peers along various characteristics leads to stronger treatment effects in the endline. An individual in Treatments 2 or 3 (i.e., the paired treatment arms) is more similar to their peer if they are in the same income group, age group, caste, marital status, or education. We construct an index combining these similarity variables.[27] Table E.1 in the appendix shows these results, where we compare Treatment 1 with various pair types and additionally control for similarity, which is normalized to be between 0 and 1.[28] The effect of close and more central peers on aspirations is significantly higher than other pair types ($p=0.02$), and also higher for take-up ($p=0.04$). The effect on take-up also outperforms being trained alone ($p=0.07$); for business, the difference is significant ($p=0.08$). These results also persist when we instead control for peer characteristics, such as their caste, age, education, income, and marital status. Close and central peers continue to significantly outperform other pair types on aspirations ($p=0.02$) and take-up ($p=0.05$), though differences relative to training alone are not statistically significant. These results are shown in Table E.2.

4.2.2 Popular Peers and Alternative Centrality Measures

To examine whether the "role model" effect documented in the literature (e.g., (Bernard et al. 2026)) extends to peer-based training, we study how outcomes depend on whether an individual is matched with a popular peer. We define popular peers as those with centrality greater than or equal to the 75th percentile of that centrality measure in the respective village. As shown in Table C.2, relative to other pair types, being paired with a popular and close peer raises aspirations — significantly so using degree centrality ($p=0.01$) and indegree ($p=0.03$), with similar patterns for eigenvector centrality ($p=0.07$). In contrast, there are no significant differences for knowledge, intentions, steps, or take-up. A full decomposition into

[27] As before, we use the method proposed in Anderson (2008) to construct the index.

[28] This implies that similarity is equal to 1 for the most similar pairs in the sample.

four cells—Popular × Close, Not Popular × Close, Popular × Far, and Not Popular × Far (Table SA.12) confirms that Popular × Close is the only cell with consistently significant aspirations effects compared to the control group across three measures of centrality.

The same pattern holds when we use alternative definitions of centrality. Results comparing treatment effects using various centrality measures are shown in Table C.1. Using eigenvector centrality, those matched with more central and socially close peers report higher aspirations than other pair types (p=0.04). Using indegree centrality, differences in aspirations are positive but not statistically significant (p=0.13). These patterns are broadly consistent with the results obtained using degree centrality in our main analysis.

Importance of Network Variables using Random Forests

We complement the above findings by implementing a predictive exercise using random forests (Breiman 2001). For each pre-specified index, we train a random forest model for the peer treatment arms that include only peer characteristics. These characteristics include both network measures (e.g., peer degree, whether peer is more central, social distance, interaction between centrality and closeness) and demographics (peer education, income, age, marital status, and caste). To assess variable importance, we measure how much the model's prediction error increases when the values of a given predictor are randomly shuffled across observations. We then report the three predictors that reduce accuracy the most. This exercise is descriptive, but it highlights which peer traits systematically explain variation in outcomes. As seen in Panel A of Figure SA.4, we find that the network characteristics of the peer are consistently selected as the most important predictors of the knowledge, aspirations, and business indices.

4.3 Long Term Effects

We now discuss the effects on outcomes measured one year after the program. About 3% of individuals in the follow-up sample had opened new businesses after one year of training. As shown in Table 3, training alone significantly improves the business steps index by 0.49 standard deviations—a result that remains statistically significant after adjusting for multiple hypothesis testing ($q<0.05$). Training with a peer shows no effects apart from a marginal increase in the steps index relative to the pure control group ($p=0.109$). Table SA.27 shows that IV estimates yield significant effects on the steps index for both training alone ($+0.55\sigma$, $p<0.05$) and training with a peer ($+0.33\sigma$, $p<0.05$) among compliers. However, none of these effects are robust to multiple hypothesis correction (all q-values equal 1), suggesting that while LATE point estimates are larger than ITT, they are not robust to multiple testing adjustment.

We also detect positive spillover effects among individuals in the spillover group within the same village. These effects are not significantly different from those trained alone, but they are higher for the steps and mindset indices compared to those trained with a peer. To explore whether these spillovers operate through direct social connections, we use the baseline village network to measure each spillover individual's exposure to neighbors assigned to treatment. On average, 30% of a spillover individual's baseline network connections were assigned to a treatment arm. We find that spillover individuals with a higher share of network connections assigned to training with a peer take significantly more entrepreneurial steps ($p < 0.10$), while the share assigned to training alone has no effect (Table E.11). Furthermore, simply having any network connection assigned to treatment significantly improves outcomes, steps, and mindset indices among spillover individuals (Table E.12).

Spillover effects persist even after applying post-double-selection Lasso (Belloni et al. 2014) (Table SA.25) to account for any potential baseline differences in individual characteristics. While attrition between endline and follow-up does not differ by treatment status, we also use inverse probability weighting to account for attrition (Tables SA.31–SA.34 in the Supplementary Appendix). The main effect of training alone on the steps index remains significant (p<0.05), though the difference between spillover and paired treatments for the steps index weakens (p=0.128).

Table 3: Effects on long-term outcomes

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.279	0.585***	0.280	0.138
	(0.190)	(0.195)	(0.203)	(0.147)
Trained Alone	0.194	0.490***	0.173	-0.0201
	(0.137)	(0.163)	(0.188)	(0.145)
Trained with Peer	0.103	0.215	0.0110	-0.0188
	(0.115)	(0.130)	(0.171)	(0.115)
Constant	0	0	0	0
	(0.0700)	(0.0721)	(0.157)	(0.0858)
Observations	576	580	580	580
R-squared	0.007	0.028	0.010	0.003
p: Alone vs Spillover	0.642	0.639	0.482	0.318
p: Alone vs Paired	0.428	0.0799	0.123	0.991
p: Paired vs Spillover	0.272	0.0533	0.0505	0.217
q: Spillover	0.219	0.0250	0.219	0.315
q: Alone	0.335	0.0250	0.502	0.804
q: Paired	1	0.777	1	1

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

Notes: This regression treats the pure control group as the base category. Standard errors are robust and clustered at the village level. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across the four outcome indices. Treatment effects on individual index components are discussed in the Supplementary Appendix (Section M).

4.3.1 Effects of Peer Identity

The weak average effects in the peer arm likely reflect a combination of productive and unproductive matches, making it important to test for heterogeneity. To do this, as before, we decompose the effects by pair types using continuous and discrete measures of social distance and peer centrality as well as relative comparisons between own and peer centrality (see Tables 4(a)-(c)). In Table 4(a), we find that individuals trained with a peer with high social distance have significantly lower treatment effects on the outcomes index, mindset index, and business practices index; all three effects survive correction for multiple testing ($q<0.10$). Table 4(b) shows that those matched with a more central peer are worse off, in that they have a significantly lower value of the steps index relative to those trained with a less central peer; this finding also holds after multiple testing adjustment ($q<0.10$).

Combined, these results suggest that the benefits of being matched with a close peer extend to the longer term, but the effects of centrality are negative. This is in contrast to the short-term results, where we detected strong motivational effects from training with a close and central peer. While central and close peers may initially boost motivation and engagement, less central but close peers might offer more durable support, potentially due to being able to interact more often. These long-term peer heterogeneity results are robust to post-double selection Lasso (Table SA.24): social distance continues to significantly reduce outcomes and business practices ($p<0.10$), while more central peers significantly reduce the steps index ($p<0.05$). Using inverse probability weighting to account for follow-up attrition (Tables SA.32–SA.33), the negative effects of social distance on outcomes, mindset, and business practices remain significant ($p<0.10$), while more central peers continue to significantly reduce steps ($p<0.05$). We will provide more evidence on these mechanisms in Section 5.

We next examine the interaction of distance and centrality. Unlike the endline results where we detected complementarity between social closeness and network centrality, Table 4(c) shows that the point estimates for those matched with a less central but close peer exceed those of other pair types across all outcomes, though none are statistically significant. In fact, central and socially close peers have a significantly lower effect on the steps index compared to those treated alone ($p=0.04$) as seen in Table D.1. This suggests that long-term benefits (in terms of taking steps to open a business) are more strongly associated with socially close peers, and the additional 'motivational' effect of being matched with a more central peer in the endline does not translate into improvements in the longer term. We will return to these follow-up results when we estimate peer effects in Section 6.

Importance of Network Variables using Random Forests: As before, we implement random forests for the paired treatment arms, and rank peer characteristics by how much the model prediction error increases when each variable is randomly permuted. Peer characteristics include both network measures (e.g., peer degree, whether peer is more central, social distance, interaction between centrality and closeness) and demographics (peer education, income, age, marital status, and caste). As shown in Panel B of Figure SA.4, peer network characteristics consistently emerge as one of the top 3 most important predictors of follow-up outcomes, often surpassing peer demographic characteristics.

Table 4: Effects on long-term outcomes by peer type

(a) By Social Distance and Peer Centrality

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.194	0.490***	0.173	-0.0201
	(0.137)	(0.164)	(0.188)	(0.146)
Trained with Peer	0.541*	0.355	0.489	0.305
	(0.316)	(0.385)	(0.343)	(0.219)
Trained with Peer × Peer distance	-0.0980**	0.0139	-0.101*	-0.0993**
	(0.0471)	(0.0707)	(0.0529)	(0.0456)
Trained with Peer × Peer degree	-0.0173	-0.0518	-0.0226	0.0107
	(0.0271)	(0.0326)	(0.0290)	(0.0222)
Constant	0	0	0	0
	(0.0702)	(0.0723)	(0.158)	(0.0860)
Observations	564	568	568	568
R-squared	0.012	0.034	0.018	0.012
q: Alone	0.338	0.0250	0.507	0.805
q: Paired	0.305	0.305	0.305	0.305
q: Paired x Dist	0.0990	0.269	0.0990	0.0990
q: Paired x Deg	0.974	0.974	0.974	0.974

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

(b) By Social Closeness and More/Less Central

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.191	0.454***	0.164	-0.0234
	(0.135)	(0.150)	(0.181)	(0.146)
Trained with Peer	0.0167	0.324	0.00222	-0.0351
	(0.174)	(0.194)	(0.206)	(0.126)
Trained with Peer (Close)	0.413*	0.0211	0.177	0.219
	(0.217)	(0.183)	(0.248)	(0.174)
Trained with Peer (More Central)	-0.103	-0.460**	-0.0753	-0.00818
	(0.230)	(0.181)	(0.198)	(0.133)
Degree	0.00648	0.00657	0.00842	0.00346
	(0.0221)	(0.0257)	(0.0182)	(0.0251)
Constant	-0.0395	-0.0372	-0.0475	-0.0258
	(0.137)	(0.148)	(0.185)	(0.162)
Observations	566	570	570	570
R-squared	0.015	0.038	0.013	0.006
q: Alone	0.340	0.0230	0.511	0.776
q: Paired	1	0.757	1	1
q: Close	0.374	0.834	0.780	0.490
q: Central	1	0.0740	1	1

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

(c) By Social Closeness Interacted with More/Less Central

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.289	0.593***	0.288	0.148
	(0.193)	(0.196)	(0.201)	(0.149)
Trained Alone	0.189	0.448***	0.164	-0.0243
	(0.135)	(0.150)	(0.181)	(0.145)
Close x Less Central	0.362	0.267	0.248	0.111
	(0.225)	(0.295)	(0.258)	(0.176)
Other Pair Types	0.0442	0.115	-0.0227	0.0104
	(0.133)	(0.140)	(0.168)	(0.128)
Degree	0.00955	0.0166	0.00817	0.00504
	(0.0221)	(0.0235)	(0.0195)	(0.0242)
Constant	-0.0546	-0.0864	-0.0463	-0.0335
	(0.136)	(0.138)	(0.188)	(0.158)
Observations	566	570	570	570
R-squared	0.011	0.031	0.014	0.004
p: Close+Less Central vs Other	0.215	0.568	0.217	0.578
p: Close+Less Central vs T1	0.433	0.533	0.727	0.403
q: Close+Less Central	0.906	0.906	0.906	0.906
q: Other Pairs	1	1	1	1

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

Notes: This regression treats individuals in the pure control villages as the base category and includes an indicator for the spillover group and those who were intended to be paired but remained unmatched. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality. Standard errors are robust and clustered at the village level. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes.

5 Mechanisms

In this section, we disentangle channels through which different types of peers affect outcomes in the short and long term, drawing on a variety of additional survey outcomes. We first examine short-term mechanisms including joint learning during training, encouragement, risk-sharing, and pooling of contacts. We then turn to longer-term mechanisms, asking whether peers facilitate risk-sharing, sustain collaboration, and remain in contact one year later. Together, these results highlight the role of peer identity over time.

First, we asked women who attended the training in pairs to describe how pairing was useful compared to training alone. 37% of women paired report having received encouragement from their partner. Another 30% report that the training material was easier to grasp because of the pairing. About 18% say that the training helped because they shared network contacts. We now examine each mechanism in detail.

Do pairs learn better together? To evaluate this mechanism, we test how the intervention affects knowledge for different types of peers. Column 1 of Table E.3 shows that while the intervention improved the knowledge index, this does not differ by whether the individual was trained alone or with a peer.

Uniquely, we also measure learning during the training itself through performance in interactive exercises. We use two additional variables measured in exercises conducted during the training: game profit and yearly profit. Game profit records performance on the first day of the training, where women played an investment and savings game. Yearly profit was measured on the last day of the training as the projected profit from a hypothetical business plan that pairs collaboratively developed. Columns 2 and 3 of Table E.3 show that these within-training measures of performance do not differ by whether the individual was trained alone or with a peer.

Women who attended the training with a socially close peer may have been better able to discuss the material being taught. Table E.4 shows that the effect of being trained with a socially close peer on training performance is significantly higher than for those in Treatment 1 or those trained with a distant peer. Specifically, close pairs earn significantly more in the business plan exercise than the control group, while far pairs show no significant effect; the difference between close and far pairs is significant (p=0.05). This suggests that socially close peers may be learning better together. The effect on projected profits does not extend to the individual game profit or the post-training knowledge index, suggesting that the benefit of social closeness may largely operate through better collaboration on joint tasks rather than individual learning.

Do pairs help each other share risk? Pairing can also improve outcomes if individuals are risk-averse and the matched peer is less risk-averse or can provide financial assistance in case business endeavors do not proceed as planned. Table SA.15 shows heterogeneous effects by the individual's own risk aversion: risk-averse individuals in paired treatments benefit significantly more from training on knowledge ($+0.52\sigma$), business index ($+0.45\sigma$), additional steps ($+0.58\sigma$), and on take-up ($+0.31\sigma$, $p<0.10$). However, risk-averse individuals in paired treatments have significantly lower aspirations (-0.36σ, $p<0.10$). For those trained alone, risk-averse individuals have significantly higher treatment effects on knowledge and steps, but not on the business index. Table SA.16 shows effects by the peer's risk aversion relative to the individual. Being paired with a less risk-averse peer does not improve outcomes and, if anything, reduces aspirations in T3 (-0.29σ, $p<0.05$).

To account for financial assistance, one year following the training, we measured the number of savings groups that the respondent was a part of and whether they had joined a new cooperative in the village. We find that those treated alone are significantly more likely to have joined a new cooperative, but the effect is not significant for those trained with a peer, suggesting that the paired treatments did not increase engagement with the wider network for saving or risk-sharing (Table SA.17). However, the lower rate of joining cooperatives among paired individuals could also be consistent with bilateral risk-sharing between training partners substituting for formal cooperative membership. If pairs provide mutual support to each other, they may have less need to seek risk-sharing through external groups. We provide additional information on network interactions later in this section.

Do pairs provide access to network connections? The third channel is gaining access to network connections by pooling contacts. Treatment 3 — in which women were paired and received a connections module — was designed precisely to test this mechanism. As previously discussed, this treatment fares better in terms of magnitude but the effects on most outcomes are not significantly different from other arms.

As noted earlier, women assigned to Treatment 3 pool 6.5 contacts on average. After controlling for social closeness and peer centrality, we find that the number of contacts pooled increases the probability of reporting a willingness to start a business with the matched peer at endline (Table E.6). At the same time, we do not find positive heterogeneous effects by the number of contacts pooled in Treatment 3 (Table E.8); the coefficient on contacts pooled $\times$ aspirations is negative (-0.55σ) but not statistically significant.

We also measure the pre-existing network overlap between pairs in this treatment arm by computing the share of common neighbors relative to the total combined unique neighbors of each pair. As shown in Table E.9, higher network overlap is associated with significantly lower treatment effects of the connections module on the knowledge index and additional steps taken ($p<0.10$), and negative but insignificant coefficients on business intentions and take-up. At the same time, the coefficient on aspirations is positive and significant ($p<0.10$) possibly because network overlap correlates with social proximity.

Considering the pair-level dyadic data, we find that among all those who attended the connections module, dyads with a higher social distance and higher gap in network degrees pool significantly more contacts, while higher network overlap has a negative albeit insignificant impact (Table E.5). While the treatment increases contact sharing along these dimensions, contact sharing remains concentrated within caste groups. As documented in Section 4, same-caste pairs share significantly more contacts than other pairs, with the effect dominating other predictors in magnitude.

We further examine whether caste homophily (as defined in Section 2) affects the number of contacts shared. Table E.7 shows that homophily is negatively and significantly associated with the number of contacts pooled among all participants who attended the connections module. Further, while same-caste pairs share significantly more contacts on average, high-homophily same-caste pairs are less likely to share than low-homophily same-caste pairs ($p=0.01$). This is consistent with decreasing returns to sameness, as there is likely to be a lot of overlap in existing links. A Lasso regression (Figure SA.5) confirms that caste similarity dominates other predictors of contact pooling, with its coefficient remaining stable under higher penalty values. Inward-looking social networks therefore limit the scope for contact sharing.

Do pairs encourage each other? A key channel driving short-term peer effects is encouragement. First, the most common benefit of pairing reported by respondents was encouragement. Regressing an indicator for whether the respondent reported encouragement as a benefit of pairing on peer characteristics in Table SA.18, we find that being matched with a more central peer increases the likelihood of reporting encouragement by 13.4 percentage points ($p<0.10$), and each additional peer network connection raises this likelihood by 3.0 percentage points ($p<0.05$).

Next, we define a popular individual as one whose centrality is greater than or equal to the 75th percentile in their village. Results in Table C.2 show that individuals paired with popular and close peers report significantly higher aspirations across all three alternative centrality measures. Aspirations are significantly higher for those matched with popular and socially close peers than other pair types. A full decomposition into four pair types (Table SA.12) confirms that Popular $\times$ Close is the only cell with consistently significant aspirations effects relative to the control group, reinforcing the complementarity between closeness and centrality for this outcome.

This is also corroborated by our main results discussed in the previous section: we detected a significant effect on aspirations and take-up for those matched with socially close peers. Moreover, even among socially close peers, business aspirations and take-up are higher when paired with more central relative to less central peers (though this difference is not statistically significant for aspirations, $p=0.13$; Table SA.11). Combined, these findings suggest that socially proximate *and* central peers can operate as motivators in the short term.

Are pairs likely to interact in the future? Finally, we find that individuals with lower social distance (i.e., socially close) to their matched peer are significantly more likely to want to open a business together and to meet in the future (Panel A of Table 5). In contrast, being assigned to a peer with high centrality does not differentially affect these outcomes. Notably, paired trainees are also significantly more likely to report wanting to open an agricultural business together (Table SA.10). This is consistent with the *parma* or *bola* system in Nepalese villages, where peers support each other with agriculture-related tasks through reciprocity-based labour provision (Messerschmidt 1981, Sherpa 2005, Bhattarai 2006). These traditional forms of collaboration may lower the perceived cost of joint agricultural enterprise for paired individuals.

One year later, we find that individuals who are trained with a peer are significantly more likely than the control group to report discussing business-related concerns with anyone in the village, though the overall rate of such discussions remains low (Table E.10). When asked whether they spoke to their matched peer, 31% report having done so. When asked what they spoke about, 83% sought advice, while the remainder discussed borrowing or lending money and setting up businesses. However, as shown in Table E.10, these network interactions were significantly more likely than in the control group when the matched peer was less central (Column 2), and within the paired arm, when the matched peer was socially close (Column 3).

Importantly, as shown in Panel B of Table 5, those matched with a same-caste peer or with someone with low social distance are significantly more likely to have interacted with each other. Being matched with someone with high centrality does not affect these outcomes. This highlights the potential role played by the ease of collaboration among individuals with a similar social standing and corroborates the previous empirical findings. Consistent with treatment effects diffusing through village networks, we find that spillover individuals with a higher share of network connections assigned to training with a peer take significantly more entrepreneurial steps, while connections to those assigned to train alone show no effect (Table E.11). Having any network connection assigned to treatment also significantly increases the outcomes, steps, and mindset indices (Table E.12).

5.1 Summary of Mechanisms

Overall, our evidence highlights two key channels through which peer identity affects outcomes. First, socially close peers facilitate collaboration: close pairs perform significantly better on joint training exercises, are more likely to plan future meetings and express willingness to start businesses together, and when they share the same caste, are significantly more likely to maintain contact one year later.

Second, central peers motivate through encouragement: respondents paired with central

Table 5: Effect of pair type on collaboration and communication

(a) Willingness to Start a Business Together

VARIABLES	(1) Pairs will meet in Future	(2) Pairs will start Business Together	(3) Pairs will meet in Future	(4) Pairs will start Business Together
Peer Centrality	-0.00227	0.00349		
	(0.00720)	(0.0137)		
Social Distance	-0.0295**	-0.0278		
	(0.0132)	(0.0198)		
Paired (High Centrality)			-0.0402	0.00969
			(0.0475)	(0.0688)
Paired (Close)			0.0822**	0.146**
			(0.0415)	(0.0677)
Constant	0.977***	0.478***	0.836***	0.285***
	(0.0619)	(0.113)	(0.0465)	(0.0804)
Observations	325	324	331	330
R-squared	0.642	0.164	0.644	0.177

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

(b) Talking to Matched Peer over 1 Year

VARIABLES	(1) Talk to matched peer	(2) Talk to matched peer	(3) Talk to matched peer	(4) Talk to matched peer
Peer Centrality	0.00472	0.00365		
	(0.0197)	(0.0194)		
Social Distance	-0.0465**	-0.0393*		
	(0.0235)	(0.0233)		
Same caste		0.247***		0.282***
		(0.0835)		(0.0801)
Paired (High Centrality)			-0.00122	-0.00736
			(0.0943)	(0.0909)
Paired (Close)			0.0853	0.0660
			(0.0961)	(0.0938)
Constant	0.472***	0.246	0.283**	0.0869
	(0.166)	(0.180)	(0.110)	(0.117)
Observations	177	177	181	181
R-squared	0.160	0.193	0.136	0.179

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

Notes: Panel (a) restricts the endline sample to participants assigned to the paired treatment arms. Panel (b) restricts the follow-up sample to the paired treatment arms. Robust standard errors in parentheses.

peers were significantly more likely to report encouragement as the reason pairing was beneficial, and this translates into higher aspirations when the central peer is also socially close. However, these short-term motivational effects fade over time. Those matched with less central peers were significantly more likely to talk to others about opening a business one year later and produce larger long-term gains on the steps index, consistent with central peers being less available for sustained interaction.

Taken together, the results suggest that ease of collaboration and motivation drive effects in the short term and interaction costs—shaped by social distance and caste boundaries—determine who is effective in the long run.

6 Estimation of Peer Effects

We present a simple theoretical framework in the appendix (Section F) in which a peer's network position affects entrepreneurial effort/intent through two competing channels: motivational benefits that increase with centrality and social closeness, and collaboration costs that rise with centrality but decline with closeness. The model yields three predictions: (i) closeness raises entrepreneurial effort (Proposition 1), (ii) peer centrality increases entrepreneurial effort only when its motivational benefit exceeds the associated collaboration cost (Proposition 2), and (iii) the effect of centrality can reverse over time as interaction costs become more salient after the training ends (Proposition 3). We now test these predictions using the experimental data.

6.1 Correlational Evidence

We use data from treated individuals in the paired treatment arms to test whether peer effects might differ based on network positions.[29] To do this, we first construct an outcome index in each wave. We capture short-term effects by constructing a standardized Anderson Index (Anderson 2008) combining all the endline variables that enter the knowledge, business aspirations, business, steps, and take-up indices. Similarly, we combine all follow-up variables into a single outcome index. We interpret these outcome indices as reduced-form proxies for entrepreneurial intent. Using these measures, we estimate the following equation:

$$e_i^* = \theta + \theta_o \mathbf{X}_i + \theta_p \mathbf{X}_j + \theta_e e_j^* + \theta_d c_{ij} e_j^* + \theta_c \phi_j e_j^*$$

where e_i^* is the outcomes index of individual i.[30] The vector $\mathbf{X}_i$ includes individual characteristics such as age, education, caste, income, and number of social connections, and $\mathbf{X}_j$ includes the analogous peer characteristics. We estimate the equation for the endline and follow-up waves separately because the empirical analysis has previously shown that centrality plays a different role depending on the time horizon over which effects are measured.

While random pairing addresses endogenous selection into pairs, the lack of exogenous variation in peer outcomes does not allow us to solve the "reflection problem" arising from the simultaneous effect of peer effort on own effort and vice versa (Manski 1993). We first implement correlations to check whether the signs align with the reduced-form results we find in our experiment.

The results are shown in Table 6.[31] We detect heterogeneity in the magnitude of peer effects depending on how central the peers are – being matched with well-connected peers improves endline outcomes. This effect persists in the endline even when we control for additional own and peer characteristics, including their age, caste, income, and education level. However,

[29] Throughout this section, we only include treated individuals. We include non-compliers or non-randomized participants who were treated with a peer, as the estimation of peer effects only relies on random pairing.

[30] We denote outcomes index as e_i^* as it is a proxy for entrepreneurial effort and is an equilibrium object derived in Section F in the appendix.

[31] The sample for the paired treatments in the follow-up is smaller, as we only include cases where we observe both the individual and their peer after 1 year.

Table 6: Effect of closeness and centrality in the short and long run

VARIABLES	(1) Outcome (Endline)	(2) Outcome (Endline)	(3) Outcome (Follow-up)	(4) Outcome (Follow-up)
Peer Outcome	-0.0287	-0.0915	0.131	0.157
	(0.104)	(0.0951)	(0.172)	(0.170)
Close x Peer Outcome	-0.116	-0.0905	0.117	0.0751
	(0.0872)	(0.0760)	(0.153)	(0.158)
Peer Degree X Peer Outcome	0.0426*	0.0474**	-0.00431	-0.00515
	(0.0220)	(0.0187)	(0.0269)	(0.0264)
Constant	0.199***	1.047***	-0.000422	0.936
	(0.0424)	(0.272)	(0.0696)	(0.706)
Observations	454	452	186	186
R-squared	0.027	0.212	0.030	0.178

Notes: This table shows the effect of the matched peer's outcome on the individual's own outcome. The first two columns use data from the endline wave, while the last two columns use data from the follow-up wave. Columns (2) and (4) additionally control for individual and peer characteristics.

these effects vanish in the follow-up wave. This pattern is consistent with our proposed framework: centrality can improve outcomes in the short run when the motivational benefit dominates, but the effect reverses in the long run as collaboration costs become more salient (Propositions 2 and 3 of the model presented in Section F).

6.2 Estimation of Peer Effects

Next, as shown in Table 7, we leverage the data collected during the two waves and regress each individual's follow-up outcome on their peer's endline outcome. This addresses the reflection problem (Manski 1993), which is otherwise not addressed solely by random matching. We find that peer centrality significantly reduces the association between peer endline outcomes and own follow-up outcomes. For peers with few connections, the estimated peer effect is large and positive (0.35–0.38 SD). Each additional peer connection reduces this effect by 0.06–0.07 SD. At the sample mean peer degree of approximately 4.6, the estimated peer effect is close to zero and statistically insignificant. Figure 2 plots the marginal peer effect at each level of peer degree showing how being paired with a more connected peer can adversely affect longer-term outcomes.

We interpret these patterns as consistent with our reduced-form findings and the predictions of our theoretical framework (Section F) rather than as structural estimates of peer effects, since correlated shocks at the pair or network-position level could still remain. While the training content was entirely homogeneous across pairs, we cannot rule out that shared exposure to common shocks during the training may contribute to the observed association as they can affect both the individual's own outcomes and the outcomes of their peer. Without additional exogenous variation in one of the peers' effort, we cannot address this concern further.

Table 7: Peer effect estimates

VARIABLES	(1) Outcome (Follow-up)	(2) Outcome (Follow-up)
Peer Outcome (Endline)	0.353**	0.376**
	(0.169)	(0.183)
Close X Peer Outcome	-0.129	-0.172
	(0.176)	(0.179)
Peer Degree X Peer Outcome	-0.0672**	-0.0644**
	(0.0315)	(0.0313)
Own Outcome (Endline)	0.130**	0.124*
	(0.0648)	(0.0707)
Constant	1.001*	0.483
	(0.565)	(0.645)
Observations	288	288
R-squared	0.141	0.278

Notes: The table reports the effect of the matched peer's endline outcome on the individual's own follow-up outcome. Close × Peer Outcome interacts the peer outcome with an indicator for whether the peer was socially close. Peer Degree × Peer Outcome interacts the peer outcome with the peer's degree centrality. Both columns control for the individual's and peer's age, income, education, caste, and network degree. Column (2) additionally includes village fixed effects.

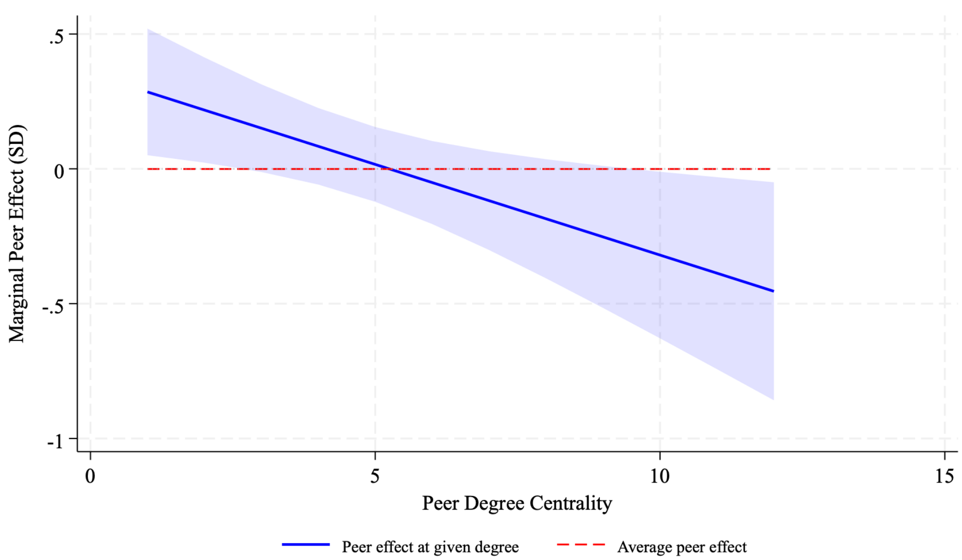


Figure 2: Marginal peer effect by peer degree centrality

Notes: The figure plots the estimated marginal effect of a 1 SD increase in peer endline outcome on own follow-up outcome, at each level of peer degree centrality, based on specification (1) of Table 7. The shaded area represents 90% confidence intervals. The dashed red line shows the sample-average peer effect, computed by averaging the marginal effect across the actual distribution of peer degrees.

Finally, it is possible that the peer's endline outcome is endogenous if it is correlated with their follow-up outcome and also affects the individual's follow-up outcome. This seems unlikely since the majority of pairs do not remain in contact. However, Table SA.30 shows that the results remain very similar even when we control for the peer's own follow-up outcome in the above regression. The negative association between being matched with a highly connected peer and the individual's follow-up outcome persists.

6.3 Feasibility of Peer Pairings

While pairings based on social distance can be easy to implement, the feasibility of pairing based on network centralities and the potential returns from such pairings remain unclear. First, it is not clear if it is feasible to collect network data and engineer such pairs in the first place. We find that 63% of individuals in our setting guess the number of connections of their randomly matched peer within a range of ± 2 connections (≈ 1 standard deviation), and about 41% guess it correctly within a range of ± 1 connection. This aligns with findings in the literature showing that individuals can possess accurate knowledge about who is central in their network (Banerjee et al. 2019), and it may not require additional data collection.

Second, it is not clear how many such pairs we can form within a village, given the limited number of 'central' individuals in a fixed village network. To check this, we simulate 10,000 counterfactual reassignments of peer pairings within each village and find that strategic pairings based on network position are indeed feasible.

For each village, we first grouped individuals into randomly chosen dyads. Then, we computed the differences in degree centralities of each dyad, and computed the average degree centrality gap in the village. We then repeat this process 10,000 times for each village. The distribution of the average centrality gaps is plotted in Figure SA.8.

The figure shows that while some villages have tightly clustered gaps, the majority display much wider distributions, implying that pairs with large gaps in centralities are indeed feasible to implement. These simulations confirm that strategic matching within villages, based on network position, is feasible and, in many cases, could be used to generate substantial treatment effects in the short-term.

7 Conclusion

Development programs frequently incorporate peer interactions, yet the peer's identity is rarely treated as a design choice. In this paper, we show that who individuals are paired with matters substantially for outcomes, and that the optimal peer depends on the time horizon. Socially close peers generate benefits in both the short and long term — they collaborate better during training, are more likely to remain in contact, and are more likely to discuss business ideas a year later. Central peers, by contrast, provide short-term motivational gains, particularly when socially close, but these gains do not translate into sustained interactions.

For entrepreneurship programs in developing countries, where starting a business requires both initial motivation and ongoing support, this distinction is consequential: the peer who inspires action today may not be the one who supports over time. Our estimates of peer effects confirm that this heterogeneity is large enough to reverse the sign of the average effect, suggesting that programs treating peers as interchangeable may miss both productive and counterproductive matches.

We also test whether peers can serve as bridges to the broader social network — a channel that development interventions increasingly aim to activate. Despite structural conditions that strongly favor bridging — 73% of pairs share no friends in common, and contact sharing increases with social distance — we find no significant impact on outcomes. Contact sharing is concentrated among individuals of the same caste, and even within same-caste pairs, those in castes with higher homophily share fewer contacts. This suggests that in settings where social networks are organized along group boundaries, facilitating introductions through peers may not be sufficient to generate new productive ties. Future interventions may need to address the underlying reluctance to engage across social boundaries, rather than simply reducing the cost of reaching new contacts.

Finally, our findings suggest that strategic peer assignment, i.e., pairing individuals based on network position, is both feasible and potentially high-return. The majority of individuals in our sample correctly estimate their peer's number of connections within one standard deviation, implying that network data collection may not be necessary. Simulations of counterfactual reassignments indicate substantial scope for design improvements. While our results demonstrate that such strategic pairings are feasible and may not require additional network data, they do not capture the private costs of networking. These costs may be higher when interacting with central peers than with less central but socially close ones. In effect, our findings speak to the net benefit of interacting with different types of peers, but the decomposition into private benefits and costs remains unknown. Costs can range from informational (e.g., knowing who is central, which we show is not binding in our setting) to interactional (e.g., confidence, time). While we cannot measure such costs, reducing them through appropriate platforms may help leverage the benefits of strategic networking.

Appendix

A Baseline Results

Table A.1: Summary statistics

Age	37.98	(10.85)
Divorced	0.00141	(0.0375)
Married	0.918	(0.274)
Unmarried	0.0669	(0.250)
Widow	0.0134	(0.115)
Higher Education (Class 11, 12)	0.104	(0.305)
Informal Education	0.126	(0.332)
No Education	0.326	(0.469)
Primary (Class 1-5)	0.155	(0.362)
Secondary (Class 6-10)	0.249	(0.432)
University	0.0402	(0.196)
Belongs to Upper Caste	0.341	(0.474)
Degree Centrality	4.585	(2.140)
Eigen Vector Centrality	0.00978	(0.0133)
Own Non Agr. Business	0.220	(0.415)
Feel not Capable	0.277	(0.448)
Willing to Open a Business	0.419	(0.494)
Risk Aversion (1-6)	4.610	(1.406)
Have no skills	0.496	(0.500)
Financial Reasons	0.239	(0.427)
No support from family	0.0129	(0.113)
Aspirations (Agricultural Expenditure)	245802.6	(331604.3)
Aspirations (Non Agri. Expenditure)	443675.3	(861565.3)
Aspirations (Income)	141669.0	(658774.2)
Aspires to Higher Income	0.890	(0.313)
Aspires to Higher Non Agri. Exp	0.229	(0.420)
Observations	2840	

Notes: This table reports summary statistics (i.e. mean and standard deviation) for baseline characteristics. The sample includes all baseline respondents with non-missing data.

Figure A.1: Experiment Design

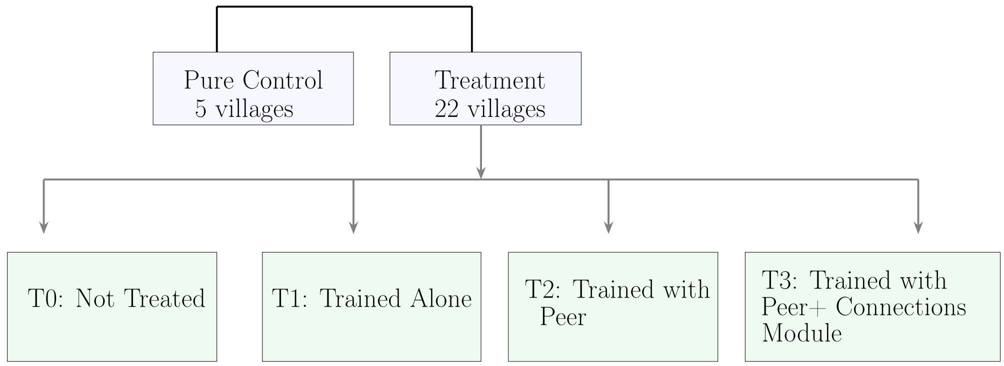


Figure A.2: The figure illustrates the random assignment of 27 villages into pure control and treatment groups. Within treatment villages, individuals were further randomized into one of three arms: trained alone (T1), trained with a peer (T2), or trained with a peer plus a connections module (T3).

B Implementation

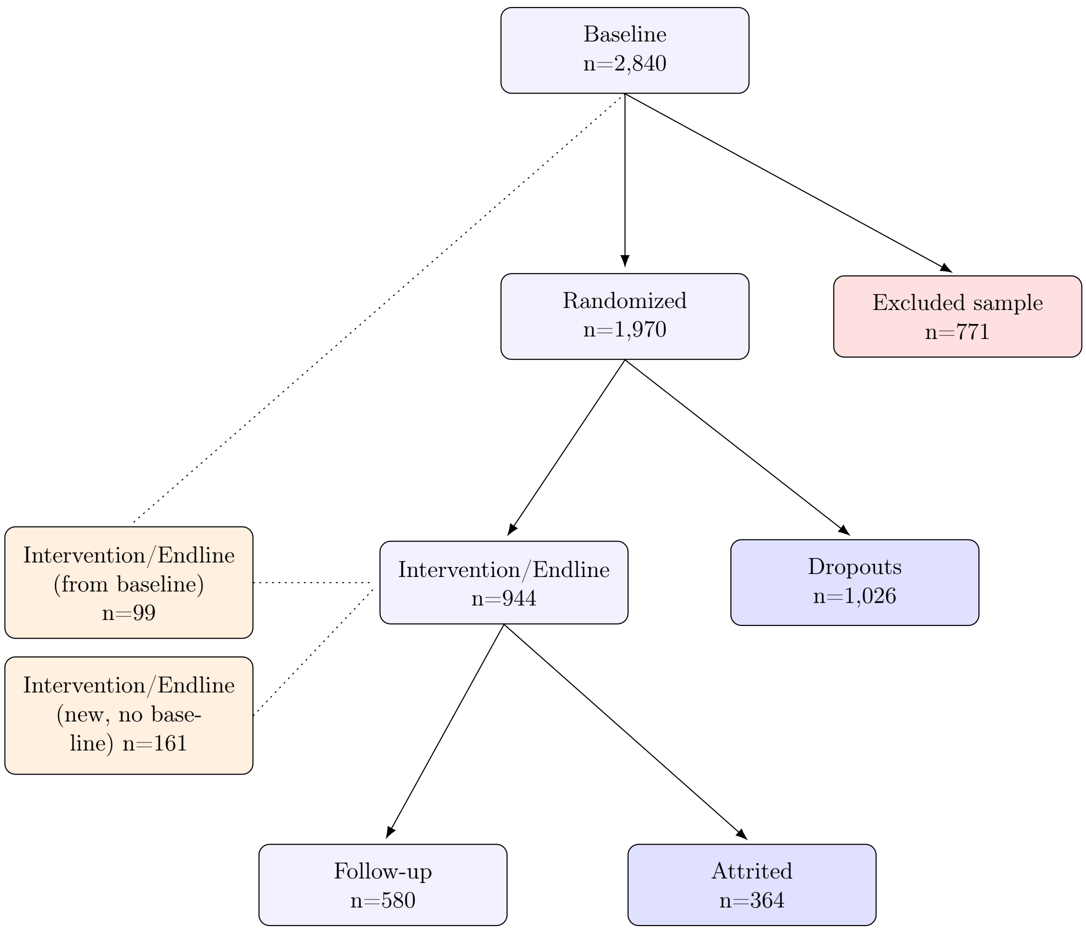


Figure B.1: The figure shows participant numbers at each stage of data collection. From the baseline sample (n = 2,840), 1,970 women were randomized into treatment arms, while 771 were excluded ex-ante. Of those randomized, 944 participated in the intervention and endline survey, and 580 were successfully followed up one year later. The "Intervention/Endline" box also includes 99 baseline respondents who attended despite not being randomized and 161 new participants without baseline data. These non-randomized individuals are excluded from all intent-to-treat analyses. An additional 260 non-randomized women, including those from two excluded villages, also completed the endline. They are excluded from the reduced form analysis as we implement intent-to-treat analysis but their partners (who were present in the randomized list) are retained.

B.1 Compliance

Compliance is defined as a binary indicator equal to one if the individual's realized treatment status matched their assignment. Overall compliance in the endline sample is 68%, or 58% when excluding pure control villages. This includes those whom we could not find and who do not appear in the endline survey (i.e., dropouts), as well as those who did not adhere to their assigned treatment, e.g. did not get trained despite being assigned to training.

Compliance is not correlated with treatment status in the endline sample (Column 1 in Table B.1) but this pattern disappears if we exclude the pure control villages (Column 2). In that case, compliance is higher for those assigned to treatment, i.e., those invited for training were more likely to attend training compared to compliance among those asked not to attend. In treated villages, 82% of those who arrived at the venue and were assigned to training (under any of the treatments) were actually trained, while the remainder were in the actual control group. Compliance is much lower for the control group as some within-village controls were included in the training to fill spots left by non-attending treatment-assigned individuals. We do not find any key differences in the characteristics of compliers and non-compliers in this subsample, as shown in Table B.2, except that compliers are more likely to be married.

To address any issues arising due to compliance, we present intention-to-treat (ITT) estimates in the main results. We also present robustness checks using comparisons with pure control villages and LATE specifications that instrument actual treatment status with treatment assignment.

Table B.1: Correlation of compliance with treatment status (endline sample)

VARIABLES	(1) Compliance	(2) Compliance
Trained Alone	-0.0562	0.301***
	(0.0929)	(0.0501)
Trained with Peer	-0.105	0.252***
	(0.0932)	(0.0511)
Trained with Peer + CM	-0.0982	0.259***
	(0.0955)	(0.0528)
Constant	0.732***	0.375***
	(0.0849)	(0.0366)
Observations	944	710
R-squared	0.010	0.058
Trained Alone vs Paired	0.276	0.327
Paired vs Paired+CM	0.868	0.901
Trained Alone vs Paired+CM	0.508	0.413

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

Notes: Compliance is defined as a binary measure indicating adherence to the assigned treatment in the endline sample. Column 2 excludes the pure control villages.

Table B.2: Baseline characteristics of compliers in treated villages (endline sample)

	(1) 0	(2) 1	(3) (1) vs. (2), p-value
Age	38.649	38.849	0.798
Income	24459.091	24311.509	0.940
No Education	0.352	0.366	0.712
Informal Education	0.154	0.156	0.950
Primary Education	0.141	0.176	0.209
Secondary Education	0.265	0.239	0.432
Higher Education	0.081	0.051	0.127
University Education	0.007	0.012	0.447
Upper Caste	0.331	0.324	0.834
Married	0.920	0.954	0.071
Divorced	0.000	0.000	
Widow	0.023	0.012	0.275
Degree Centrality	5.144	5.359	0.233
Willing to Open a Business	0.452	0.404	0.212
Risk Aversion (1-6)	4.525	4.781	0.016

Notes: This table compares baseline characteristics of individuals who complied with the assigned treatment with those who did not, in treated villages, excluding dropouts (i.e. who are not present in the endline survey).

B.2 Initial Dropouts

Table B.3 examines whether dropouts from our randomized sample (N=1970), i.e., those who did not appear in the endline survey (N=944), differ by treatment status. In the full sample (Column 1), dropout rates are higher among those assigned to treatment compared to the pooled control group, which includes both controls in treated villages and households in pure control villages. This difference is driven by survey logistics rather than participant selection as women in treated villages completed the survey at the training centre, whereas those in pure control villages were surveyed at home. We present various robustness checks to ensure that any results in the follow up survey (where we use the data on the pure control group) are not affected by this selection.

Once we restrict attention to treated villages (Columns 2–3), this pattern disappears: dropout rates are not systematically different between those assigned to treatment and those in the control group within the same village. This suggests that the higher attrition observed in Column 1 reflects survey modality rather than selection related to treatment assignment i.e., surveys were largely conducted at training centers in treatment villages but at households in pure control villages. Importantly, survey attrition does not differ between those trained alone and those trained with a peer, which is the main comparison of interest. While dropout is somewhat higher in the Peer + Connections Module arm relative to the Alone and Peer arms, the differences are only marginally significant, and participants were not informed ex-ante about the module.

Table B.3: Initial dropouts by treatment status

VARIABLES	(1) Initial Dropout	(2) Initial Dropout	(3) Initial Dropout
Trained Alone	0.0694*	-0.0276	-0.0276
	(0.0386)	(0.0351)	(0.0351)
Trained with Peer	0.0757*	-0.0213	
	(0.0411)	(0.0351)	
Trained with Peer + CM	0.133***	0.0364	
	(0.0398)	(0.0349)	
Paired			0.00748
			(0.0304)
Constant	0.464***	0.561***	0.561***
	(0.0359)	(0.0248)	(0.0248)
Observations	1,970	1,606	1,606
R-squared	0.010	0.003	0.001
p-value: Alone vs. Peer	0.830	0.858	
p-value: Peer vs. Peer+CM	0.0291	0.0991	
p-value: Alone vs. Peer+CM	0.0488	0.0673	
p-value: Alone vs. Paired			0.249

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

Notes: This table compares the proportion of dropouts by intended treatment status within the randomized sample where Column 2 excludes the pure control villages and Column 3 combines the paired treatment arms into one variable.

Table B.4 further compares baseline characteristics of individuals who dropped out before the endline to those who remained. Differences are limited to education and the number of connections in the social network. The endline sample may therefore exclude socially isolated women, which could affect external validity. However, since dropouts cannot be related to treatment assignment, treatment effects are unbiased. Other characteristics, including age, income, caste, risk aversion, and baseline willingness to open a business, are balanced. Baseline balance checks between the treated and control groups (Table B.6) show that these initial dropouts do not result in systematic selection across key variables, including baseline willingness to open a business, which is in fact, higher in the control group.

Table B.4: Baseline characteristics of dropouts

	(1) 0	(2) 1	(3) (1) vs. (2), p-value
Age	38.765	37.970	0.145
Income	24373.855	25470.248	0.422
No Education	0.360	0.401	0.097
Informal Education	0.155	0.094	0.000
Primary Education	0.161	0.142	0.287
Secondary Education	0.250	0.249	0.959
Higher Education	0.064	0.092	0.036
University Education	0.010	0.023	0.031
Upper Caste	0.327	0.324	0.908
Married	0.939	0.913	0.042
Divorced	0.000	0.002	0.157
Widow	0.017	0.015	0.703
Degree Centrality	5.268	4.245	0.000
Willing to Open a Business	0.424	0.422	0.931
Risk Aversion (1-6)	4.673	4.676	0.972

Notes: This table compares baseline characteristics of individuals who dropped out before the endline with those who remained in the sample.

B.3 Follow-up Attrition

Table B.5 tests whether attrition in the follow-up survey differs by treatment assignment. Attrition is 40% overall (constant) and does not vary significantly across Spillover, Trained Alone, or Trained with Peer. Pairwise tests confirm no differential attrition. This reduces concern that long-run results are driven by selection.

Table B.5: Correlation of follow-up survey attrition with treatment status

VARIABLES	(1) Attrition
Spillover	0.00170
	(0.0511)
Trained alone	-0.0506
	(0.0559)
Trained with Peer	-0.0173
	(0.0560)
Constant	0.402***
	(0.0372)
Observations	944
R-squared	0.002
Spillover==T1	0.368
T1=Paired	0.623
Spillover=Paired	0.752

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

Notes: Attrition is a binary variable equal to 1 if the individual is observed in the endline survey but not in the follow up. This regression compares attrition rates among individuals in the spillover, T1, and paired treatment arms with those in the pure control group. Standard errors are robust and clustered at the village level.

B.4 Balance Tables

Table B.6: Balance test for within-village comparisons (endline sample)

	(1) 0	(2) 1	(3) 2	(4) 3	(5) (1) vs. (2), p-value	(6) (1) vs. (3), p-value	(7) (1) vs. (4), p-value	(8) (2) vs. (3), p-value	(9) (2) vs. (4), p-value	(10) (3) vs. (4), p-value
Willing to Open Business	0.46	0.39	0.40	0.45	0.17	0.21	0.85	0.90	0.25	0.31
Income	24666.47	25317.20	25653.33	23907.65	0.81	0.72	0.79	0.91	0.65	0.58
Income Source- Agri.	0.85	0.87	0.84	0.91	0.48	0.82	0.06	0.35	0.22	0.03
Income Source- Business	0.06	0.06	0.10	0.10	0.87	0.17	0.22	0.12	0.16	0.92
Income Source- Job	0.03	0.02	0.01	0.03	0.66	0.23	0.89	0.42	0.57	0.20
Income Source- Remit.	0.01	0.01	0.01	0.02	0.53	0.96	0.59	0.56	0.27	0.55
Income Source- Other	0.07	0.03	0.04	0.01	0.08	0.14	0.00	0.76	0.21	0.13
Age	40.37	39.70	38.45	39.96	0.53	0.08	0.71	0.23	0.81	0.17
Elementary Education	0.10	0.15	0.19	0.16	0.18	0.02	0.11	0.30	0.75	0.50
Higher Education	0.06	0.09	0.05	0.07	0.29	0.91	0.66	0.24	0.56	0.58
Informal Education	0.32	0.31	0.30	0.33	0.84	0.67	0.83	0.81	0.68	0.53
University Education	0.02	0.01	0.01	0.01	0.29	0.30	0.35	0.99	0.91	0.92
Secondary Education	0.32	0.23	0.27	0.21	0.06	0.32	0.03	0.36	0.69	0.20
Degree	4.79	5.28	5.29	5.32	0.04	0.04	0.03	0.96	0.88	0.93
Brahmin	0.05	0.05	0.07	0.03	0.89	0.45	0.35	0.36	0.42	0.09
Chhetri	0.31	0.30	0.23	0.27	0.94	0.09	0.50	0.10	0.54	0.32
Dalit	0.02	0.01	0.04	0.01	0.29	0.14	0.35	0.02	0.91	0.02
Newar	0.39	0.41	0.50	0.45	0.66	0.04	0.26	0.11	0.47	0.42
Janjati	0.23	0.23	0.16	0.24	0.92	0.09	0.95	0.11	0.87	0.09

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The balance tests compare characteristics of the control group within treated villages (0) with those in Treatment 1, Treatment 2, and Treatment 3 for the endline sample. The p-values in Columns 5-10 indicate if the difference is statistically significant.

Table B.7: Balance test for the follow-up sample

	(1) 0	(2) 1	(3) 2	(4) 3	(5) (1) vs. (2), p-value	(6) (1) vs. (3), p-value	(7) (1) vs. (4), p-value	(8) (2) vs. (3), p-value	(9) (2) vs. (4), p-value	(10) (3) vs. (4), p-value
Income	25545.029	27694.207	25364.717	23304.280	0.585	0.946	0.507	0.612	0.410	0.396
Income Source- Agri.	0.865	0.869	0.842	0.968	0.934	0.601	0.003	0.672	0.010	0.033
Income Source- Business	0.024	0.000	0.017	0.000	0.123	0.689	0.123	0.171	0.171	
Income Source- Job	0.037	0.008	0.017	0.032	0.054	0.216	0.754	0.569	0.122	0.465
Income Source- Remit.	0.012	0.008	0.017	0.022	0.718	0.698	0.586	0.569	0.255	0.804
Income Source- Other	0.057	0.025	0.042	0.011	0.191	0.491	0.019	0.387	0.520	0.114
Age	39.861	39.459	38.975	38.430	0.781	0.501	0.361	0.626	0.307	0.650
Elementary Education	0.155	0.180	0.200	0.161	0.512	0.285	0.898	0.659	0.738	0.479
Higher Education	0.078	0.090	0.058	0.086	0.749	0.587	0.830	0.333	0.884	0.314
Informal Education	0.314	0.270	0.325	0.333	0.483	0.875	0.789	0.471	0.346	0.916
University Education	0.008	0.008	0.008	0.011	0.997	0.987	0.833	0.991	0.854	0.860
Secondary Education	0.314	0.221	0.242	0.269	0.050	0.190	0.458	0.650	0.265	0.648
Degree	4.807	5.438	5.350	5.226	0.009	0.036	0.045	0.743	0.466	0.594
Brahmin	0.073	0.041	0.050	0.022	0.531	0.610	0.233	0.757	0.559	0.160
Chhetri	0.257	0.311	0.225	0.301	0.616	0.760	0.700	0.102	0.857	0.146
Newar	0.196	0.238	0.158	0.237	0.586	0.604	0.661	0.052	0.986	0.264
Dalit	0.424	0.402	0.542	0.430	0.871	0.440	0.970	0.054	0.668	0.070
Tamang	0.049	0.008	0.025	0.011	0.264	0.553	0.283	0.535	0.858	0.327

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

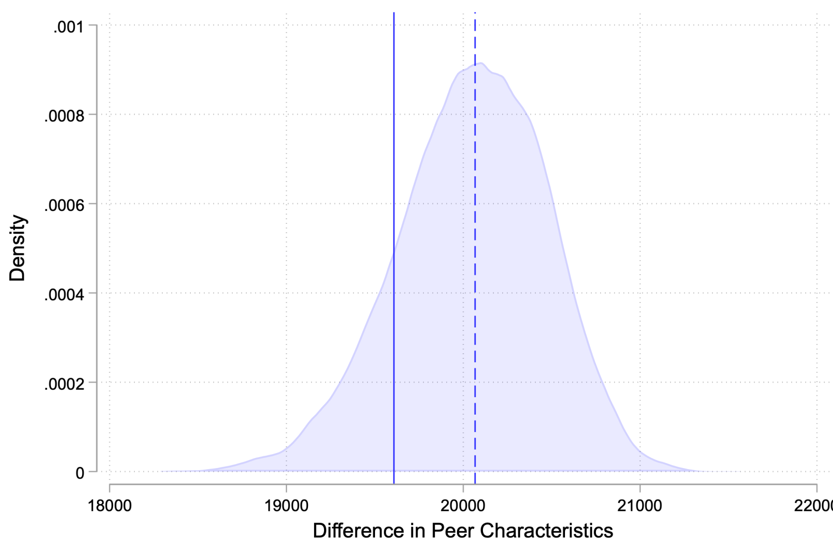
Notes: The balance tests compare characteristics of those in the pure control group (0) with those in the spillover group(1), Treatment 1 (2), and paired treatment arms (3) for the follow-up sample. The p-values in Columns 5-10 indicate if the difference is statistically significant.

B.5 Randomization Inference for Peer Assignment

We test whether the peer pairs created during the intervention are systematically more or less similar than would be expected under random assignment. To do this, we construct a statistic that captures the overall difference between peers across a set of characteristics including income, source of income, age, education level, caste, marital status, willingness to start a business, income aspirations, agricultural expenditure aspirations, overall degree, outdegree, and risk aversion. In each iteration of the simulation, we match individuals within each village into dyads, compute the sum of the quadratic differences in their baseline characteristics, and add this across the entire sample.

Comparing the observed statistic to the distribution obtained from the simulations allows us to assess whether actual peers are unusually similar or dissimilar. The distribution and the statistic from the actual assignment is plotted in Figure B.2, indicating that the observed pairings are not systematically different from what random assignment would generate.

Figure B.2: Randomization inference on peer assignment



Notes: The histogram shows the distribution of peer similarity under 10,000 random re-assignments. The solid vertical line marks the statistic computed from actual peer pairs and the dashed vertical line plots the median difference computed from the re-assignments.

C Endline Results

Table C.1: Impact of training using alternative centrality measures (Close × more central)

(a) Indegree

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.088	0.103	0.083	0.114	0.211**
	(0.094)	(0.114)	(0.102)	(0.099)	(0.095)
Close x More Central	0.311**	0.541	0.212	0.048	0.225
	(0.155)	(0.342)	(0.183)	(0.189)	(0.154)
Other Pair Types	0.180**	0.024	0.066	0.090	0.228**
	(0.088)	(0.109)	(0.100)	(0.097)	(0.095)
Observations	700	700	697	678	684
R-squared	0.059	0.010	0.023	0.022	0.032
p: Close+Central = Other	0.365	0.131	0.420	0.820	0.984
p: Close+Central = T1	0.131	0.204	0.476	0.720	0.924

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

(b) Eigenvector

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.105	0.107	0.097	0.117	0.216**
	(0.094)	(0.112)	(0.101)	(0.098)	(0.094)
Close x More Central	0.223*	0.555*	0.196	0.108	0.154
	(0.134)	(0.287)	(0.154)	(0.152)	(0.138)
Other Pair Types	0.216**	-0.019	0.079	0.085	0.253***
	(0.088)	(0.104)	(0.101)	(0.097)	(0.095)
Observations	700	700	697	678	684
R-squared	0.048	0.018	0.019	0.024	0.033
p: Close+Central = Other	0.955	0.0441	0.437	0.876	0.445
p: Close+Central = T1	0.346	0.121	0.510	0.950	0.631

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

Notes: Each panel uses a different centrality measure (degree centrality results are in Table 2(c) of the main text). "Close × More Central" = paired with a socially close peer who is more central than the respondent. Coefficients are relative to within-village controls. Robust standard errors in parentheses.

Table C.2: Impact of training using alternative definition of peer popularity (Popular × close)

(a) Degree

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.086	0.113	0.080	0.113	0.209**
	(0.094)	(0.114)	(0.101)	(0.099)	(0.094)
Popular x Close	0.144	0.591**	0.123	0.084	0.318***
	(0.125)	(0.264)	(0.136)	(0.135)	(0.105)
Other Pair Types	0.208**	-0.051	0.066	0.083	0.196**
	(0.089)	(0.103)	(0.103)	(0.100)	(0.099)
Observations	700	700	697	678	684
R-squared	0.061	0.022	0.027	0.023	0.035
p: Popular+Close = Other	0.566	0.0134	0.674	0.990	0.214
p: Popular+Close = T1	0.615	0.0702	0.748	0.824	0.248

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

(b) Indegree

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.088	0.108	0.084	0.114	0.212**
	(0.094)	(0.114)	(0.102)	(0.099)	(0.095)
Popular x Close	0.199	0.564**	0.115	0.017	0.347***
	(0.128)	(0.280)	(0.146)	(0.151)	(0.109)
Other Pair Types	0.195**	-0.030	0.077	0.102	0.197**
	(0.089)	(0.104)	(0.102)	(0.099)	(0.099)
Observations	700	700	697	678	684
R-squared	0.058	0.018	0.023	0.022	0.034
p: Popular+Close = Other	0.975	0.0322	0.793	0.557	0.139
p: Popular+Close = T1	0.347	0.102	0.829	0.504	0.172

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

(c) Eigenvector

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.104	0.110	0.096	0.117	0.215**
	(0.094)	(0.112)	(0.101)	(0.098)	(0.094)
Popular x Close	0.102	0.457*	-0.026	-0.024	0.164
	(0.132)	(0.256)	(0.153)	(0.150)	(0.135)
Other Pair Types	0.248***	-0.011	0.135	0.120	0.253***
	(0.089)	(0.106)	(0.102)	(0.098)	(0.096)
Observations	700	700	697	678	684
R-squared	0.049	0.014	0.020	0.025	0.033
p: Popular+Close = Other	0.224	0.0685	0.291	0.316	0.499
p: Popular+Close = T1	0.983	0.181	0.419	0.332	0.691

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

Notes: Each panel uses a different centrality measure. "Popular × Close" = paired with a peer above the 75th percentile in centrality who is also socially close. Coefficients are relative to within-village controls. Robust standard errors in parentheses.

D Follow-up Results

Table D.1: Effects on long-term outcomes by type of peer

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.291	0.594***	0.289	0.148
	(0.192)	(0.196)	(0.201)	(0.149)
Trained Alone	0.185	0.447***	0.161	-0.0258
	(0.135)	(0.149)	(0.182)	(0.145)
Close x More Central	0.416	-0.0326	0.00385	0.278
	(0.516)	(0.188)	(0.405)	(0.242)
Other Pair Types	0.0548	0.167	0.0285	-0.00822
	(0.127)	(0.155)	(0.171)	(0.117)
Degree	0.0165	0.0186	0.0129	0.00782
	(0.0211)	(0.0246)	(0.0184)	(0.0238)
Constant	-0.0888	-0.0964	-0.0696	-0.0472
	(0.132)	(0.142)	(0.185)	(0.156)
Observations	566	570	570	570
R-squared	0.011	0.031	0.012	0.006
p: Close x More Central vs Other Pairs	0.523	0.328	0.951	0.199
p: Close x More Central vs T1	0.676	0.0439	0.708	0.189

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

Notes: This regression treats individuals in the pure control villages as the base category and includes an indicator for the spillover group and those who were intended to be paired but remained unmatched. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality. Standard errors are robust and clustered at the village level.

E Mechanisms

Table E.1: Impact of being trained with a central friend after controlling for similarity

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.0861	0.107	0.0794	0.112	0.207**
	(0.0936)	(0.114)	(0.101)	(0.0988)	(0.0943)
Close x More Central	0.230	0.699*	0.485**	0.211	0.566***
	(0.198)	(0.367)	(0.232)	(0.219)	(0.203)
Other Pair Types	0.124	-0.0763	0.311	0.207	0.394**
	(0.182)	(0.219)	(0.201)	(0.196)	(0.188)
Degree	0.0459***	0.0124	0.0529***	0.0246*	0.0312**
	(0.0122)	(0.0197)	(0.0153)	(0.0146)	(0.0135)
Similarity Index	0.0942	0.0838	-0.461	-0.224	-0.353
	(0.277)	(0.356)	(0.324)	(0.306)	(0.329)
Constant	-0.220**	-0.0592	-0.253**	-0.118	-0.150
	(0.1000)	(0.114)	(0.110)	(0.107)	(0.108)
Observations	700	700	697	678	684
R-squared	0.061	0.023	0.031	0.023	0.038
p: Close+Central vs Other Pairs	0.433	0.0195	0.247	0.979	0.0408
p: Close+Central vs T1	0.454	0.108	0.0784	0.647	0.0699

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

Notes: This regression treats within-village controls as the base category. We additionally control for the "Similarity Index" which is a weighted index of similarity along characteristics including income, age, caste, marital status, and education and also control for degree-centrality of the individual. Robust standard errors in parentheses.

Table E.2: Effect of being matched with a central friend after controlling for other characteristics

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Close x More Central	0.444	0.745	0.236	-0.130	-0.0694
	(0.327)	(0.656)	(0.428)	(0.435)	(0.376)
Other Pair Types	0.367	0.0121	0.0646	-0.157	-0.236
	(0.311)	(0.535)	(0.430)	(0.402)	(0.389)
Constant	-0.222**	-0.0574	-0.253**	-0.123	-0.145
	(0.101)	(0.115)	(0.111)	(0.108)	(0.110)
Observations	699	699	696	677	683
R-squared	0.071	0.041	0.038	0.035	0.048
p: Close+Central vs Other Pairs	0.573	0.0194	0.259	0.865	0.0503
p: Close+Central vs T1	0.268	0.333	0.714	0.576	0.459

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

Notes: This regression treats within-village controls as the base category. We additionally control for peer characteristics such as peer age, caste, whether they are married, baseline income, and education. Robust standard errors in parentheses.

Table E.3: Effect on learning during and after the training

VARIABLES	(1) Knowledge Index	(2) Profit (Game)	(3) Profit (Business Plan)
Trained alone	0.114		
	(0.0944)		
Trained with Peer	0.0678	52.51	203,402
	(0.0863)	(41.13)	(263,911)
Constant	0	273.3***	534,159***
	(0.0753)	(34.06)	(168,030)
Observations	710	437	436
R-squared	0.002	0.004	0.001
T1==T2/T3	0.519		

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

Notes: This regression treats within-village controls as the base category in Column 1 and Treatment 1 as the base category in Column 2 and 3. Robust standard errors in parentheses.

Table E.4: Learning by whether pairs are friends or not

VARIABLES	(1) Knowledge Index	(2) Profit (Game)	(3) Profit (Business Plan)
Trained alone	0.114		
	(0.0946)		
Paired (Close)	0.233**	55.71	1.011e+06*
	(0.110)	(58.02)	(566,051)
Paired (Far)	0.216**	48.64	-111,855
	(0.0910)	(44.11)	(254,422)
Constant	-0	273.3***	534,159***
	(0.0754)	(34.14)	(168,425)
Observations	705	432	432
R-squared	0.045	0.005	0.023
p: Close vs Far	0.864	0.897	0.0508

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

Notes: Column 1 uses within-village controls as the base category, while Columns 2 and 3 use those treated alone as the base category. Robust standard errors in parentheses.

Table E.5: Correlation of number of contacts pooled and similarity with peer

VARIABLES	(1) Number of links pooled	(2) Number of links pooled
Same age group		-0.603
		(0.673)
Same caste		1.552***
		(0.435)
Same education		0.689
		(0.493)
Same income group		-0.267
		(0.633)
Quadratic diff. in connections	0.0149*	0.0151*
	(0.00861)	(0.00885)
Network Distance	0.176*	-0.00701
	(0.0906)	(0.116)
Network Overlap	-0.312	-0.549
	(0.467)	(0.418)
Constant	5.570***	5.235***
	(0.368)	(0.346)
Number of dyads	90	90
R-squared	0.072	0.205

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

Notes: This regression compares dyads in Treatment Arm 3, including all pairs who attended the connections module, regardless of assignment. Robust standard errors in parentheses.

Table E.6: Perceived future interactions by number of contacts shared

VARIABLES	(1) Pairs will start Business Together	(2) Pairs will meet in Future
Number of links pooled	0.0603***	-0.00580
	(0.0186)	(0.0126)
Paired (Socially Close)	0.228**	0.0726
	(0.0893)	(0.0456)
Paired (High Centrality)	-0.0473	-0.0862
	(0.0905)	(0.0685)
Constant	-0.00934	0.938***
	(0.115)	(0.0724)
Observations	157	157
R-squared	0.091	0.024

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

Notes: Columns include all individuals who attended Treatment 3, regardless of assignment. Robust standard errors in parentheses.

Table E.7: Relationship between caste homophily and number of contacts pooled

VARIABLES	(1) Number of contacts pooled	(2) Number of contacts pooled
High Homophily	-1.108**	-0.476
	(0.461)	(0.478)
Same caste	1.059***	2.569**
	(0.255)	(0.983)
High Homophily x Same Caste		-1.822*
		(1.006)
Constant	6.020***	5.556***
	(0.429)	(0.460)
Observations	134	134
R-squared	0.163	0.211
p: High vs. Low Homophily (same-caste pairs)		0.0106

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

Notes: Columns include all individuals who attended Treatment 3, regardless of assignment. Homophily is measured using the inbreeding homophily index (Currarini et al. 2009): $\frac{\text{share of same-caste links} - \text{share of own caste in population}}{1 - \text{share of own caste in population}}$, which ranges from -1 (maximum heterophily) through 0 (no excess same-type preference) to 1 (maximum homophily) and is comparable across groups of different sizes. Robust standard errors are in parentheses.

Table E.8: Heterogeneous treatment effects by number of contacts shared

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.114	0.125	0.108	0.126	0.243**
	(0.0946)	(0.113)	(0.102)	(0.0987)	(0.0970)
Treatment with Peer	0.175*	0.154	0.0669	0.0473	0.225**
	(0.0949)	(0.131)	(0.109)	(0.106)	(0.106)
Treatment with Peer + Connections Module	0.331**	0.330	0.277	0.364**	0.434**
	(0.153)	(0.227)	(0.183)	(0.176)	(0.181)
T3 X Number of contacts pooled	-0.161	-0.554	-0.256	-0.399	-0.297
	(0.225)	(0.345)	(0.279)	(0.263)	(0.304)
Constant	0	0	-0	0	0
	(0.0755)	(0.0755)	(0.0755)	(0.0759)	(0.0757)
Observations	685	685	682	665	669
R-squared	0.043	0.006	0.014	0.025	0.031

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

Notes: This regression treats within-village controls as the base category. Robust standard errors in parentheses.

Table E.9: Heterogeneous treatment effects by pre-existing network overlap (T3)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
T1 (Trained Alone)	0.0862	0.110	0.0799	0.112	0.208**
	(0.0937)	(0.114)	(0.101)	(0.0989)	(0.0944)
T2 (Peer)	0.222*	0.204	0.0400	-0.0341	-0.0295
	(0.130)	(0.209)	(0.166)	(0.152)	(0.171)
T3 (Peer + CM)	0.341***	0.00716	0.154	0.139	0.0255
	(0.129)	(0.172)	(0.179)	(0.166)	(0.179)
T3 × Network Overlap	-2.521*	7.331*	-1.680	-3.247*	-1.532
	(1.483)	(4.227)	(1.670)	(1.800)	(1.422)
Degree	0.0456***	0.00436	0.0518***	0.0262*	0.0298**
	(0.0123)	(0.0202)	(0.0153)	(0.0148)	(0.0134)
Peer Degree	-0.00973	-0.0206	-0.000373	0.0130	0.0474**
	(0.0187)	(0.0265)	(0.0241)	(0.0201)	(0.0212)
Constant	-0.218**	-0.0209	-0.248**	-0.126	-0.143
	(0.100)	(0.116)	(0.111)	(0.108)	(0.108)
Observations	681	681	678	661	666
R-squared	0.066	0.025	0.029	0.031	0.040

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

Notes: Network overlap is equal to the share of common friends of i and j divided by the total number of unique friends. The regression treats the within village control group as the base category and controls for own degree centrality and peer degree centrality. Robust standard errors in parentheses.

Table E.10: Effects on network interactions during the year after training

VARIABLES	(1) Talk to anyone about business	(2) Talk to anyone about business	(3) Talk to anyone about business
Spillover	0.0723**	0.0730**	
	(0.0290)	(0.0297)	
Trained Alone	0.0513	0.0494*	
	(0.0328)	(0.0253)	
Close x More Central		0.0911	
		(0.0718)	
Close x Less Central		0.155**	
		(0.0746)	
Far x Central		0.0636	
		(0.0424)	
Far x Less Central		0.0614*	
		(0.0372)	
Degree		0.00423	
		(0.00523)	
Trained with Peer	0.0655**		
	(0.0277)		
Peer Centrality			-0.00752
			(0.00873)
Social Distance			-0.0353***
			(0.0132)
Constant	0.0143	-0.00634	0.261***
	(0.0123)	(0.0264)	(0.0917)
Observations	579	569	201
R-squared	0.013	0.027	0.045
Paired=Nonpaired	0.689		

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

Notes: Column 1 treats the pure control group as the base category with pooled treatment indicators. Column 2 additionally decomposes the paired treatment arms into four peer-type categories. Column 3 restricts the sample to the paired treatment arms and regresses on continuous measures of peer centrality and social distance. Standard errors are robust and clustered at the village level.

Table E.11: Spillover effects by share of network connections assigned to treatment

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.161	0.359*	0.295	0.162
	(0.181)	(0.196)	(0.244)	(0.173)
Trained Alone	0.183	0.450***	0.158	-0.0268
	(0.133)	(0.150)	(0.182)	(0.145)
Trained with Peer	0.103	0.215	0.0100	-0.0138
	(0.118)	(0.132)	(0.168)	(0.114)
Spillover × Share Trained Alone	0.538	-0.440	0.333	-0.340
	(0.735)	(0.596)	(0.677)	(0.775)
Spillover × Share Trained with Peer	0.391	1.454*	-0.210	0.113
	(0.480)	(0.723)	(0.555)	(0.480)
Constant	-0.113	-0.0707	-0.0949	-0.0565
	(0.133)	(0.147)	(0.187)	(0.149)
Observations	574	578	578	578
R-squared	0.011	0.038	0.012	0.005

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

Notes: "Spillover" is an indicator for within-village controls in treated villages. "Share Trained Alone" and "Share Trained with Peer" are the shares of the individual's baseline network connections assigned to Treatment 1 and Treatment 2/3, respectively. All regressions control for an individual's network degree. Standard errors are robust and clustered at the village level.

Table E.12: Spillover exposure: any network connection assigned to treatment

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	-0.0413	0.160	0.00252	0.183
	(0.159)	(0.193)	(0.220)	(0.199)
Trained Alone	0.185	0.453***	0.161	-0.0269
	(0.133)	(0.150)	(0.182)	(0.145)
Trained with Peer	0.104	0.217	0.0116	-0.0139
	(0.118)	(0.132)	(0.167)	(0.114)
Spillover × Any Treated	0.487*	0.630**	0.418**	-0.0501
	(0.266)	(0.302)	(0.193)	(0.242)
Constant	-0.0959	-0.0465	-0.0760	-0.0573
	(0.133)	(0.152)	(0.188)	(0.150)
Observations	574	578	578	578
R-squared	0.016	0.037	0.017	0.004

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

Notes: "Spillover" is an indicator for within-village controls in treated villages. "Any Treated" is an indicator equal to one if the individual has at least one baseline network connection who was assigned to any treatment arm. All regressions control for network degree. Standard errors are robust and clustered at the village level.

F Theoretical Framework

We now present a simple model to understand the patterns in the data and estimate peer effects, exploiting the variation in the identity of the peer. The model has two ingredients: a motivation channel through which peers provide encouragement, and a collaboration cost channel that captures the difficulty of sustained interactions.

Consider two matched women i and j who choose entrepreneurial effort $e_{it} \geq 0$ at time periods $t \in \{0, 1\}$, where $t = 0$ denotes the training/endline period and $t = 1$ denotes the one–year follow-up. Utility in period t can be written as follows:

$$U_{it} = \theta_t e_{it} - \frac{a}{2} e_{it}^2 \; + \; \beta_t \, M(c_{ij}, \phi_j) \, e_{it} e_{jt} \; - \; \kappa_t \, C(c_{ij}, \phi_j) \, e_{it} e_{jt}, \tag{1}$$

where $\theta > 0$, $a > 0$, c_{ij} is a measure of closeness between i and j and ϕ_j is peer j's centrality. The parameters β_t and κ_t measure the relative importance of motivation and collaboration frictions at horizon t.

Motivation. The term $\beta_t M(c_{ij}, \phi_j) e_{it} e_{jt}$ captures the motivational benefit of working with a peer. A central peer (high ϕ_j) has more connections in the village and can serve as a role model (Bernard et al. 2026, Riley 2024), signaling that entrepreneurship is feasible. Social closeness (high c_{ij}) can amplify this motivational effect, as individuals who are socially proximate can learn better together and support each other.

Collaboration cost. The term $\kappa_t C(c_{ij}, \phi_j) e_{it} e_{jt}$ captures the cost of coordinating with the peer. This cost has a natural microfoundation: a peer with ϕ_j connections allocates their available time and attention across these relationships, so the time they can devote to i could be proportional to $1/\phi_j$, which is decreasing in centrality. The training program subsidizes the cost of interaction, keeping κ_0 low. After the training ($t = 1$), maintaining contact requires active effort and the peer's time constraint can become binding, so $\kappa_1 > \kappa_0$. Social closeness, on the other hand, lowers coordination costs as closer individuals are more likely to share social contexts and face lower barriers to meeting.

Throughout, subscripts denote partial derivatives, so that $M_c \equiv \frac{\partial M(c,\phi)}{\partial c}$, $M_\phi \equiv \frac{\partial M(c,\phi)}{\partial \phi}$, $C_c \equiv \frac{\partial C(c,\phi)}{\partial c}$, and $C_\phi \equiv \frac{\partial C(c,\phi)}{\partial \phi}$. As per our empirical findings, $M_\phi(c, \phi) \geq 0$ (more central peers are more motivating), $M_c(c, \phi) \geq 0$ (closer peers are more motivating), $C_c(c, \phi) < 0$ (closeness reduces collaboration costs), and $C_\phi(c, \phi) > 0$ (collaborating with more central peers is costlier).

F.1 Best Responses and Equilibrium.

Differentiating (1) with respect to e_{it} yields

$$\frac{\partial U_{it}}{\partial e_{it}} = \theta_t - ae_{it} + \left[\beta_t M(c_{ij}, \phi_j) - \kappa_t C(c_{ij}, \phi_j)\right] e_{jt} = 0, \tag{2}$$

so the linear best response is

$$e_{it} = A_t + b_t(\phi_j, c_{ij})\, e_{jt}, \quad A_t = \frac{\theta_t}{a}, \quad b_t(\phi_j, c_{ij}) = \frac{\beta_t M(c_{ij}, \phi_j) - \kappa_t C(c_{ij}, \phi_j)}{a}. \tag{3}$$

Lemma: Equilibrium existence and uniqueness If $a > 0$ and $b_t(\phi_i, c_{ij})\, b_t(\phi_j, c_{ij}) < 1$, then a unique interior Nash equilibrium exists and is given by:

$$e_{it} = \frac{A_t + b_t(\phi_j, c_{ij}) A_t}{1 - b_t(\phi_i, c_{ij})\, b_t(\phi_j, c_{ij})}. \tag{4}$$

This follows from solving the pair of best-response equations simultaneously.

Proposition 1: Closeness increases effort. Given $M_c(c_{ij}, \phi) \geq 0$, $C_c(c_{ij}, \phi) < 0$, and $b_t(\phi_i, c_{ij})\, b_t(\phi_j, c_{ij}) < 1$, if peer efforts are strategic complements, that is $b_t(\phi_m, c_{ij}) > 0$ for both $m \in \{i, j\}$ (equivalently $\beta_t M(c_{ij}, \phi) > \kappa_t C(c_{ij}, \phi)$), then holding (ϕ_i, ϕ_j) fixed,

$$\frac{\partial e_{it}}{\partial c_{ij}} \; > \; 0.$$

Intuitively, when the motivational effect of peer effort exceeds the cost of collaboration (i.e. $\beta_t M(c_{ij}, \phi) > \kappa_t C(c_{ij}, \phi)$), peer efforts are complements. In this case, greater closeness both raises motivation and lowers collaboration costs, and increases equilibrium effort.

Proposition 2: Peer centrality increases effort if and only if motivation is higher than collaboration cost. Given $M_\phi(c_{ij}, \phi) \geq 0$, $C_\phi(c_{ij}, \phi) > 0$, and $b_t(\phi_i, c_{ij})\, b_t(\phi_j, c_{ij}) < 1$, if peer efforts are strategic complements, that is $b_t(\phi_m, c_{ij}) > 0$ for both $m \in \{i, j\}$, then the effect of peer centrality on effort is positive if and only if the motivational effect of centrality exceeds the associated collaboration cost, i.e.

$$\frac{\partial b_t(\phi_j, c_{ij})}{\partial \phi_j} = \frac{1}{a}\Big(\beta_t M_\phi(c_{ij}, \phi_j) - \kappa_t C_\phi(c_{ij}, \phi_j)\Big) \;>\; 0.$$

Intuitively, when effort choices are complements, higher peer centrality raises equilibrium effort only when it increases motivation more than it raises collaboration costs.

The next proposition highlights how the effect of peer centrality can vary over time due to its differential effects on motivation and collaboration.

Proposition 3: The effect of peer centrality can reverse over time If peer efforts are strategic complements, that is $b_t(\phi_m, c_{ij}) > 0$ for both $m \in \{i, j\}$, then higher peer centrality can increase effort in the short run but reduce it in the long run.

This is because at $t = 0$, when κ_0 is small enough, the motivational effect of centrality dominates and $\frac{\partial e_{i0}}{\partial \phi_j} > 0$. At $t = 1$, a larger κ_1 increases the relative cost of collaboration, and for sufficiently high κ_1 the effect can turn negative, so that $\frac{\partial e_{i1}}{\partial \phi_j} < 0$. Intuitively, while the motivational benefits of working with a central peer are immediate, the costs of coordination with a central (but likely busy) peer are larger in the longer term.

References

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.

Asiedu, E., Lambon-Quayefio, M., Truffa, F. & Wong, A. (2023), 'Female entrepreneurship and professional networks', *PEDL research paper* .

Banerjee, A., Breza, E., Chandrasekhar, A. G., Duflo, E., Jackson, M. O. & Kinnan, C. (2024), 'Changes in social network structure in response to exposure to formal credit markets', *Review of Economic Studies* **91**(3), 1331–1372.

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

Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2019), 'Using gossips to spread information: Theory and evidence from two randomized controlled trials', *The Review of Economic Studies* **86**(6), 2453–2490.

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

Beaman, L., Duflo, E., Pande, R. & Topalova, P. (2012), 'Female leadership raises aspirations and educational attainment for girls: A policy experiment in india', *science* **335**(6068), 582–586.

Belloni, A., Chernozhukov, V. & 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. & Yekutieli, D. (2006), 'Adaptive linear step-up procedures that control the false discovery rate', *Biometrika* **93**(3), 491–507.

Bernard, T., Dercon, S., Orkin, K., Schinaia, G. & Taffesse, A. S. (2026), 'The future in mind: Aspirations and long-term outcomes in rural ethiopia', *The Quarterly Journal of Economics* p. qjag002.

Bernard, T. & Seyoum Taffesse, A. (2014), 'Aspirations: An approach to measurement with validation using ethiopian data', *Journal of African economies* **23**(2), 189–224.

Bhattarai, S. (2006), The bola or parma of the newar in manamaiju village. the significance of a farm labor exchange system among indigenous peasants in nepal., Master's thesis, Universitetet i Tromsø.

Boucher, V., Rendall, M., Ushchev, P. & Zenou, Y. (2024), 'Toward a general theory of peer effects', *Econometrica* **92**(2), 543–565.

Bramoullé, Y., Djebbari, H. & Fortin, B. (2020), 'Peer effects in networks: A survey', *Annual Review of Economics* **12**(1), 603–629.

Breiman, L. (2001), 'Random forests', *Machine learning* **45**(1), 5–32.

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

Breza, E. & Chandrasekhar, A. G. (2019), 'Social networks, reputation, and commitment:

evidence from a savings monitors experiment', *Econometrica* **87**(1), 175–216.

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

Brooks, W., Donovan, K. & Johnson, T. R. (2018), 'Mentors or teachers? microenterprise training in kenya', *American Economic Journal: Applied Economics* **10**(4), 196–221.

Cai, J., Janvry, A. D. & Sadoulet, E. (2015), 'Social networks and the decision to insure', *American Economic Journal: Applied Economics* **7**(2), 81–108.

Cai, J. & Szeidl, A. (2018), 'Interfirm relationships and business performance', *The Quarterly Journal of Economics* **133**(3), 1229–1282.

Calvó-Armengol, A., Patacchini, E. & Zenou, Y. (2009), 'Peer effects and social networks in education', *The review of economic studies* **76**(4), 1239–1267.

Campos, F. M. L., Coleman, R. D., Conconi, A., Donald, A. A., Gassier, M., Goldstein, M. P., Chavez, Z. L., Mikulski, J., Milazzo, A., Paryavi, M., Pierotti, R. S., O'Sullivan, M. B. & Vaillant, J. (2019), 'Profiting from parity: Unlocking the potential of women's businesses in africa : Main report'.
URL: *http://documents.worldbank.org/curated/en/501971553025918098/Main-Report*

Carranza, E., Dhakal, C. & Love, I. (2018), 'Female entrepreneurs: How and why are they different?'.

Chandrasekhar, A. G., Kinnan, C. & Larreguy, H. (2018), 'Social networks as contract enforcement: Evidence from a lab experiment in the field', *American Economic Journal: Applied Economics* **10**(4), 43–78.

Currarini, S., Jackson, M. O. & Pin, P. (2009), 'An economic model of friendship: Homophily, minorities, and segregation', *Econometrica* **77**(4), 1003–1045.

De Mel, S., McKenzie, D. & Woodruff, C. (2014), 'Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from sri lanka', *Journal of Development Economics* **106**, 199–210.

Duflo, E., Dupas, P. & Kremer, M. (2011), 'Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya', *American economic review* **101**(5), 1739–1774.

Durlauf, S. N. & Fafchamps, M. (2005), Social capital, *in* P. Aghion & S. N. Durlauf, eds, 'Handbook of Economic Growth', Vol. 1B, Elsevier, Amsterdam, pp. 1639–1699.

Fafchamps, M. & Quinn, S. (2018), 'Networks and manufacturing firms in africa: Results from a randomized field experiment', *The World Bank Economic Review* **32**(3), 656–675.

Field, E., Jayachandran, S. & Pande, R. (2010), 'Do traditional institutions constrain female entrepreneurship? a field experiment on business training in india', *American Economic Review* **100**(2), 125–129.

Field, E., Jayachandran, S., Pande, R. & Rigol, N. (2016), 'Friendship at work: Can peer effects catalyze female entrepreneurship?', *American Economic Journal: Economic Policy* **8**(2), 125–153.

Glaeser, E. L., Laibson, D. & Sacerdote, B. (2002), 'An economic approach to social capital', *The economic journal* **112**(483), F437–F458.

Granovetter, M. S. (1973), 'The strength of weak ties', *American journal of sociology*

78(6), 1360–1380.

Lerner, J. & Malmendier, U. (2013), 'With a little help from my (random) friends: Success and failure in post-business school entrepreneurship', *The Review of Financial Studies* **26**(10), 2411–2452.

Macours, K. & Vakis, R. (2014), 'Changing households' investment behaviour through social interactions with local leaders: Evidence from a randomised transfer programme', *The Economic Journal* **124**(576), 607–633.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *The review of economic studies* **60**(3), 531–542.

McKenzie, D. & Woodruff, C. (2014), 'What are we learning from business training and entrepreneurship evaluations around the developing world?', *The World Bank Research Observer* **29**(1), 48–82.

McKenzie, D., Woodruff, C., Bjorvatn, K., Bruhn, M., Cai, J., Gonzalez-Uribe, J., Quinn, S., Sonobe, T. & Valdivia, M. (2021), 'Training entrepreneurs', *VoxDevLit* **1**(2), 3.

Messerschmidt, D. A. (1981), '" nogar" and other traditional forms of cooperation in nepal: Significance for development', *Human organization* pp. 40–47.

Putnam, R. D. (2000), *Bowling alone: The collapse and revival of American community*, Simon and schuster.

Riley, E. (2024), 'Role models in movies: The impact of queen of katwe on students' educational attainment', *Review of Economics and Statistics* **106**(2), 334–351.

Sacerdote, B. (2011), Peer effects in education: How might they work, how big are they and how much do we know thus far?, *in* 'Handbook of the Economics of Education', Vol. 3, Elsevier, pp. 249–277.

Sherpa, N. S. (2005), Indigenous peoples of nepal and traditional knowledge, *in* 'International workshop on Traditional Knowledge', pp. 21–23.

Vasilaky, K. N. & Leonard, K. L. (2018), 'As good as the networks they keep? improving outcomes through weak ties in rural uganda', *Economic Development and Cultural Change* **66**(4), 755–792.

Vega-Redondo, F., Pin, P., Ubfal, D., Benedetti-Fasil, C., Brummitt, C., Domínguez, M., Rubera, G., Hovy, D. & Fornaciari, T. (2023), 'Networking entrepreneurs', *Working Paper.* .

Zárate, R. A. (2023), 'Uncovering peer effects in social and academic skills', *American Economic Journal: Applied Economics* **15**(3), 35–79.

Supplementary (Online) Appendix

Contents

A Baseline Descriptives

Table SA.1: Correlations between number of network connections and baseline characteristics

	Degree Centrality
Age	-0.00957
Divorced	0.0469*
Married	0.0800***
Unmarried	-0.0830***
Widow	-0.0257
Higher Education (Class 11, 12)	-0.0402*
Informal Education	0.0404*
No Education	-0.0744***
Primary (Class 1-5)	0.0216
Secondary (Class 6-10)	0.0847***
University	-0.0552**
Belongs to Upper Caste	0.0719***
Own Non Agr. Business	-0.0437*
Willing to Open a Business	0.0171
Risk Aversion (1-6)	-0.0681***
Aspirations (Agricultural Expenditure)	0.0236
Aspirations (Non Agri. Expenditure)	-0.0374
Aspirations (Income)	-0.0856***
Aspires to Higher Income	-0.0382
Aspires to Higher Non Agri. Exp	-0.0218

Notes: This table reports variable-wise correlations between degree-centrality (i.e. number of connections in the social network) and other baseline characteristics.

Table SA.2: Correlations between whether or not an individual has opened a business and their baseline characteristics

	Own Non Agr. Business
Age	-0.0944***
Divorced	0.00270
Married	-0.0354
Unmarried	0.0499**
Widow	-0.0250
Higher Education (Class 11, 12)	0.129***
Informal Education	-0.00791
No Education	-0.181***
Primary (Class 1-5)	0.00828
Secondary (Class 6-10)	0.00354
University	0.221***
Belongs to Upper Caste	-0.00378
Degree Centrality	-0.0437*
Eigen Vector Centrality	-0.0491**
Risk Aversion (1-6)	-0.0910***
Aspirations (Agricultural Expenditure)	-0.0364
Aspirations (Non Agri. Expenditure)	0.273***
Aspirations (Income)	0.155***
Aspires to Higher Income	-0.0429*
Aspires to Higher Non Agri. Exp	0.382***

Notes: This table reports variable-wise correlations between whether an individual has already opened a non-agricultural business and other baseline characteristics.

Table SA.3: Correlations between whether or not an individual is willing to open a business and their baseline characteristics

	Willing to open a business
Age	-0.391***
Divorced	0.0187
Married	-0.0508*
Unmarried	0.0965***
Widow	-0.0826***
Higher Education (Class 11, 12)	0.167***
Informal Education	-0.104***
No Education	-0.300***
Primary (Class 1-5)	0.0367
Secondary (Class 6-10)	0.251***
University	0.100***
Belongs to Upper Caste	-0.0541*
Degree Centrality	0.0171
Eigen Vector Centrality	0.00140
Own Non Agr. Business	0
Feel not Capable	-0.00718
Willing to Open a Business	1
Risk Aversion (1-6)	-0.103***
Aspirations (Agricultural Expenditure)	0.0455*
Aspirations (Non Agri. Expenditure)	0.167***
Aspirations (Income)	-0.0250
Aspires to Higher Income	0.0372
Aspires to Higher Non Agri. Exp	0.239***

Notes: This table reports variable-wise correlations between baseline willingness to open a non-agricultural business and other baseline characteristics.

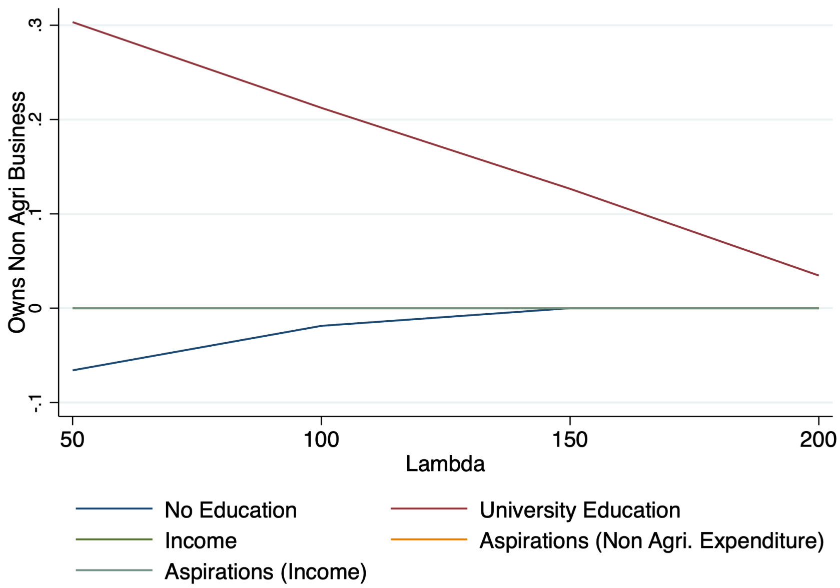


Figure SA.1: This figure plots correlations between selected variables and owning a business for different values of penalty parameters in a Lasso regression. Only variables selected out of a large list of demographic and network variables at $\lambda = 100$ are displayed. No Education and University Education are binary variables equal to 1 if the condition is true, and High Non. Agricultural Expenditure Aspirations is a binary variable equal to 1 if aspirations are higher than the baseline level.

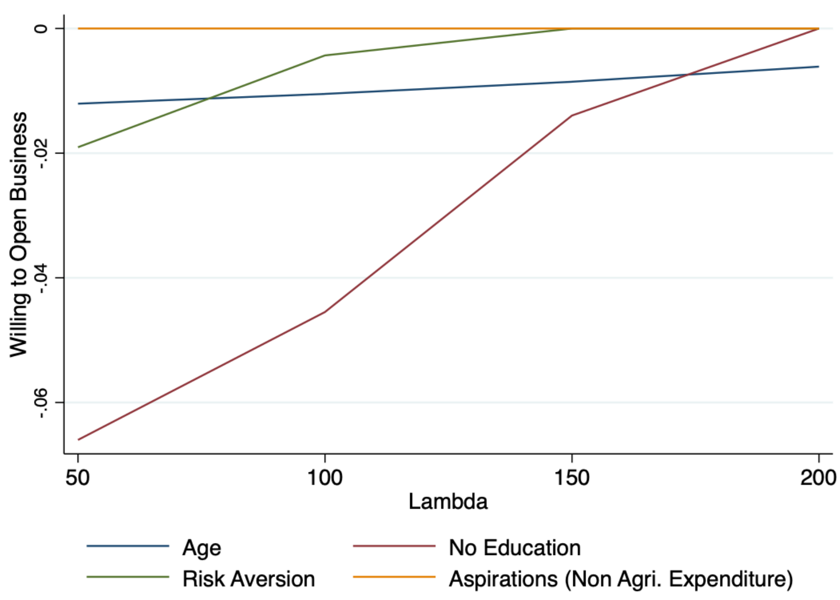


Figure SA.2: This figure plots correlations between selected variables and willingness to open businesses for different values of penalty parameters in a Lasso regression. Only variables selected out of a large list of demographic and network variables at $\lambda = 100$ are displayed. No Education is a binary variable equal to 1 if true. Risk Aversion is measured on a scale of 1-6 using a series of choices between lotteries and fixed outcomes presented to participants in the baseline survey.

B Balance and Design

Table SA.4: Balance by social closeness within paired arms (endline sample)

	(1) 0	(2) 1	(3) (1) vs. (2), p-value
Willing to Open Business	0.420	0.472	0.437
Income (Winsorized)	20988.616	37308.989	0.010
Age	39.665	37.562	0.092
Elementary Education	0.134	0.225	0.082
Secondary Education	0.244	0.270	0.657
Higher Education	0.061	0.067	0.844
Informal Education	0.354	0.258	0.112
University Education	0.006	0.011	0.688
Brahmin	0.061	0.079	0.606
Chhetri	0.262	0.292	0.615
Dalit	0.012	0.022	0.568
Newar	0.470	0.461	0.893
Janjati/Tamang	0.195	0.146	0.315
Income: Agriculture	0.854	0.899	0.287
Income: Business	0.006	0.034	0.172
Income: Job	0.024	0.022	0.923
Income: Remittance	0.000	0.000	
Income: Other	0.024	0.000	0.045

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample restricted to individuals in paired treatment arms (T2 and T3) who were matched with a peer. Close peers are defined as those with network distance ≤ 2. P-values from t-tests with standard errors clustered at the village level.

Table SA.5: Balance by relative centrality within paired arms (endline sample)

	(1) 0	(2) 1	(3) (1) vs. (2), p-value
Willing to Open Business	0.439	0.442	0.965
Income (Winsorized)	23439.691	31802.062	0.152
Age	38.454	39.887	0.263
Elementary Education	0.151	0.186	0.487
Secondary Education	0.263	0.247	0.782
Higher Education	0.092	0.010	0.002
Informal Education	0.296	0.361	0.293
University Education	0.007	0.010	0.760
Brahmin	0.072	0.062	0.746
Chhetri	0.270	0.289	0.747
Dalit	0.026	0.000	0.045
Newar	0.454	0.485	0.639
Janjati/Tamang	0.178	0.165	0.796
Income: Agriculture	0.888	0.845	0.342
Income: Business	0.013	0.021	0.665
Income: Job	0.033	0.010	0.206
Income: Remittance	0.000	0.000	
Income: Other	0.000	0.031	0.081

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample restricted to individuals in paired treatment arms (T2 and T3) who were matched with a peer. More central peers are those with strictly higher degree centrality than the respondent. P-values from t-tests with standard errors clustered at the village level.

Table SA.6: Balance by social closeness within paired arms (follow-up sample)

	(1) 0	(2) 1	(3) (1) vs. (2), p-value
Willing to Open Business	0.447	0.426	0.811
Income (Winsorized)	21241.972	31946.809	0.077
Age	39.583	37.234	0.152
Elementary Education	0.167	0.191	0.716
Secondary Education	0.250	0.298	0.546
Higher Education	0.056	0.106	0.315
Informal Education	0.361	0.277	0.295
University Education	0.009	0.021	0.604
Brahmin	0.037	0.064	0.507
Chhetri	0.278	0.277	0.988
Dalit	0.500	0.489	0.904
Newar	0.185	0.170	0.823
Janjati/Tamang	0.000	0.000	
Income: Agriculture	0.898	0.894	0.933
Income: Business	0.000	0.021	0.317
Income: Job	0.019	0.043	0.459
Income: Remittance	0.000	0.000	
Income: Other	0.028	0.000	0.083

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample restricted to individuals in paired treatment arms (T2 and T3) who were matched with a peer and appear in the follow-up survey. Close peers are defined as those with network distance ≤ 2. P-values from t-tests with standard errors clustered at the village level.

Table SA.7: Balance by relative centrality within paired arms (follow-up sample)

	(1) 0	(2) 1	(3) (1) vs. (2), p-value
Willing to Open Business	0.455	0.424	0.714
Income (Winsorized)	24536.626	24130.000	0.933
Age	38.044	40.483	0.126
Elementary Education	0.209	0.117	0.126
Secondary Education	0.275	0.267	0.914
Higher Education	0.099	0.017	0.022
Informal Education	0.297	0.400	0.198
University Education	0.011	0.017	0.776
Brahmin	0.066	0.017	0.114
Chhetri	0.275	0.300	0.740
Dalit	0.495	0.500	0.948
Newar	0.165	0.183	0.772
Janjati/Tamang	0.000	0.000	
Income: Agriculture	0.923	0.867	0.283
Income: Business	0.000	0.017	0.318
Income: Job	0.033	0.017	0.518
Income: Remittance	0.000	0.000	
Income: Other	0.000	0.033	0.155

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample restricted to individuals in paired treatment arms (T2 and T3) who were matched with a peer and appear in the follow-up survey. More central peers are those with strictly higher degree centrality than the respondent. P-values from t-tests with standard errors clustered at the village level.

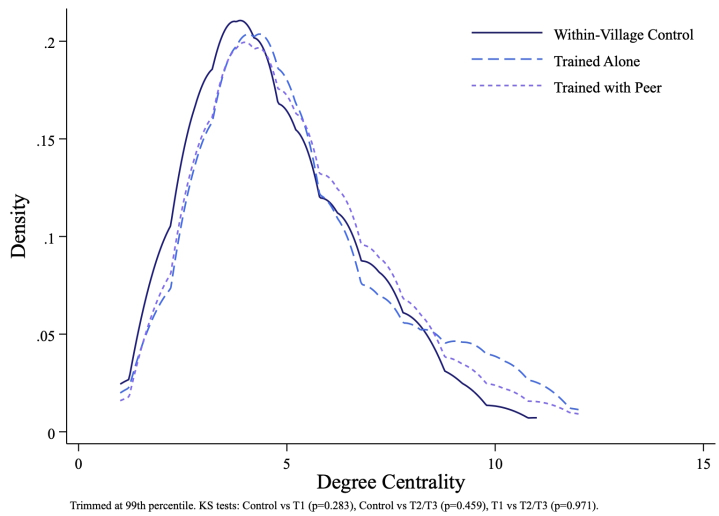
Table SA.8: Balance test for the endline sample (with pure control group)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	0	1	2	3	(1) vs. (2), p-value	(1) vs. (3), p-value	(1) vs. (4), p-value	(2) vs. (3), p-value	(2) vs. (4), p-value	(3) vs. (4), p-value
Willing to Open Business	0.43	0.39	0.40	0.45	0.57	0.64	0.72	0.91	0.27	0.35
Income	24094.72	25317.20	25653.33	23907.65	0.74	0.63	0.95	0.92	0.69	0.56
Income Source- Agri.	0.92	0.87	0.84	0.91	0.27	0.12	0.88	0.49	0.18	0.07
Income Source- Business	0.11	0.06	0.10	0.10	0.31	0.88	0.85	0.20	0.18	0.91
Income Source- Job	0.02	0.02	0.01	0.03	0.99	0.23	0.64	0.33	0.66	0.36
Income Source- Remit.	0.01	0.01	0.01	0.02	0.67	0.80	0.37	0.56	0.11	0.54
Income Source- Other	0.04	0.03	0.04	0.01	0.65	0.83	0.16	0.76	0.27	0.07
Age	39.12	39.70	38.45	39.96	0.73	0.67	0.63	0.24	0.75	0.16
Elementary Education	0.18	0.15	0.19	0.16	0.36	0.67	0.68	0.33	0.70	0.49
Higher Education	0.09	0.09	0.05	0.07	0.99	0.37	0.66	0.29	0.43	0.54
Informal Education	0.32	0.31	0.30	0.33	0.89	0.77	0.82	0.81	0.63	0.54
University Education	0.01	0.01	0.01	0.01	0.74	0.75	0.82	0.99	0.92	0.92
Secondary Education	0.28	0.23	0.27	0.21	0.24	0.79	0.15	0.19	0.55	0.13
Degree	4.78	5.28	5.29	5.32	0.07	0.10	0.02	0.96	0.86	0.90
Brahmin	0.06	0.05	0.07	0.03	0.85	0.87	0.63	0.29	0.62	0.10
Chhetri	0.23	0.30	0.23	0.27	0.66	1.00	0.78	0.07	0.48	0.22
Dalit	0.06	0.01	0.04	0.01	0.21	0.72	0.22	0.31	0.92	0.23
Newar	0.46	0.41	0.50	0.45	0.86	0.87	0.99	0.16	0.39	0.42
Janjati	0.19	0.23	0.16	0.24	0.79	0.80	0.74	0.09	0.88	0.11

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The balance tests compare characteristics of the pure control group (0) with those in Treatment 1, Treatment 2, and Treatment 3 for the endline sample. The p-values in Columns 5-10 indicate if the difference is statistically significant.

Figure SA.3: Distribution of degree centrality by treatment group



Notes: Kernel density estimates of degree centrality for controls within treated villages, Trained Alone (T1), and Trained with Peer (T2 and T3 pooled). Kolmogorov-Smirnov tests confirm no statistically significant differences in distributions across groups. Sample trimmed at 99th percentile.

Table SA.9: Correlation between endline and follow-up outcomes

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Practices Index
Knowledge Index	0.0577	-0.00254	0.0998*	0.000946
	(0.0567)	(0.0620)	(0.0520)	(0.0497)
Aspirations Index	0.280***	0.196***	0.176***	0.0709**
	(0.0830)	(0.0613)	(0.0550)	(0.0344)
Business Index	-0.0750	0.0957	0.0134	0.0793
	(0.0840)	(0.0734)	(0.0714)	(0.0671)
Additional Steps	0.156**	-0.00401	0.148**	0.0642
	(0.0696)	(0.0651)	(0.0623)	(0.0633)
Take-up Index	-0.1000*	0.144**	0.0712	0.0110
	(0.0567)	(0.0697)	(0.0575)	(0.0524)
Observations	561	565	565	565
R-squared	0.089	0.071	0.098	0.028

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

Notes: Each column reports coefficients from regressing the follow up outcome (column header) on endline outcomes. Endline outcomes are self-reported measures collected immediately after training. Follow-up outcomes are measured one year later and capture actual behaviors. Sample restricted to individuals appearing in both surveys. Robust standard errors in parentheses.

C Additional Endline Tables

Table SA.10: Types of businesses that individuals are willing to open

VARIABLES	(1) Agricultural Business	(2) Sewing	(3) Shop/Parlor	(4) Other Business
Trained alone	0.0571	0.0532	-0.121*	0.0106
	(0.0702)	(0.0475)	(0.0633)	(0.0300)
Trained with Peer	0.137**	-0.00753	-0.123**	-0.00617
	(0.0627)	(0.0389)	(0.0580)	(0.0252)
Constant	0.505***	0.108***	0.344***	0.0430**
	(0.0520)	(0.0322)	(0.0495)	(0.0211)
Observations	395	395	395	395
R-squared	0.013	0.007	0.014	0.001

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

Notes: This regression treats within-village controls as the base category. Business type is equal to 1 if the individual reports wanting to open such a business and reports being ready to invest. Robust standard errors in parentheses.

Table SA.11: Effects on short-term outcomes by type of peer (all)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.0855	0.111	0.0794	0.112	0.204**
	(0.0937)	(0.114)	(0.101)	(0.0989)	(0.0943)
Close x More Central	0.281*	0.743**	0.236	0.0900	0.376***
	(0.146)	(0.332)	(0.155)	(0.165)	(0.0941)
Close x Less Central	0.0827	0.181	-0.0222	0.0406	0.0524
	(0.131)	(0.201)	(0.159)	(0.151)	(0.158)
Far x Central	0.158	-0.182	0.00490	0.107	0.288**
	(0.119)	(0.121)	(0.146)	(0.137)	(0.131)
Far x Non Central	0.234**	-0.0288	0.113	0.0856	0.204*
	(0.0965)	(0.129)	(0.125)	(0.118)	(0.119)
Degree	0.0471***	0.00277	0.0528***	0.0263	0.0375***
	(0.0129)	(0.0198)	(0.0163)	(0.0160)	(0.0142)
Constant	-0.225**	-0.0133	-0.253**	-0.126	-0.180
	(0.103)	(0.114)	(0.114)	(0.112)	(0.111)
Observations	700	700	697	678	684
R-squared	0.063	0.026	0.029	0.023	0.038
p: Central and Close vs Less Central and Close	0.234	0.128	0.191	0.802	0.0357
p: Central and Close vs Central and Distant	0.431	0.00658	0.209	0.927	0.462
p: Central and Close vs Less Central and Distant	0.741	0.0228	0.468	0.979	0.119
p: Socially Close vs Distant	0.884	0.00433	0.704	0.799	0.750
p: Less Central vs More Central	0.551	0.302	0.567	0.781	0.0637

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

Notes: This regression treats within-village controls as the base category. Robust standard errors in parentheses.

Table SA.12: Full decomposition: Popular/Not Popular × Close/Far

(a) Degree

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.086	0.114	0.080	0.112	0.208**
	(0.094)	(0.114)	(0.101)	(0.099)	(0.094)
Popular x Close	0.144	0.593**	0.122	0.084	0.317***
	(0.125)	(0.264)	(0.136)	(0.135)	(0.105)
Not Popular x Close	0.216	0.122	0.030	0.025	-0.014
	(0.159)	(0.216)	(0.200)	(0.193)	(0.198)
Popular x Far	0.171	-0.141	-0.025	0.089	0.334***
	(0.118)	(0.120)	(0.140)	(0.135)	(0.113)
Not Popular x Far	0.233**	-0.042	0.144	0.097	0.169
	(0.095)	(0.127)	(0.127)	(0.118)	(0.125)
Observations	700	700	697	678	684
R-squared	0.061	0.023	0.028	0.023	0.040
p: Popular x Close = Not Popular x Close	0.676	0.141	0.671	0.778	0.0953
p: Popular= Not Popular	0.512	0.286	0.776	0.842	0.0368
p: Close = Far	0.831	0.0105	0.906	0.762	0.397

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

(b) Indegree

VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.088	0.109	0.084	0.114	0.211**
	(0.094)	(0.114)	(0.102)	(0.099)	(0.095)
Popular x Close	0.200	0.567**	0.115	0.017	0.344***
	(0.128)	(0.280)	(0.146)	(0.152)	(0.109)
Not Popular x Close	0.154	0.215	0.095	0.139	0.034
	(0.150)	(0.225)	(0.174)	(0.156)	(0.171)
Popular x Far	0.137	-0.135	0.065	0.110	0.351***
	(0.112)	(0.125)	(0.146)	(0.129)	(0.114)
Not Popular x Far	0.255***	-0.052	0.078	0.081	0.158
	(0.098)	(0.127)	(0.125)	(0.121)	(0.125)
Observations	700	700	697	678	684
R-squared	0.059	0.022	0.023	0.023	0.038
p: Popular x Close = Not Popular x Close	0.785	0.310	0.922	0.514	0.0733
p: Popular= Not Popular	0.716	0.468	0.979	0.694	0.0206
p: Close = Far	0.850	0.00789	0.795	0.882	0.543

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

(c) Eigenvector

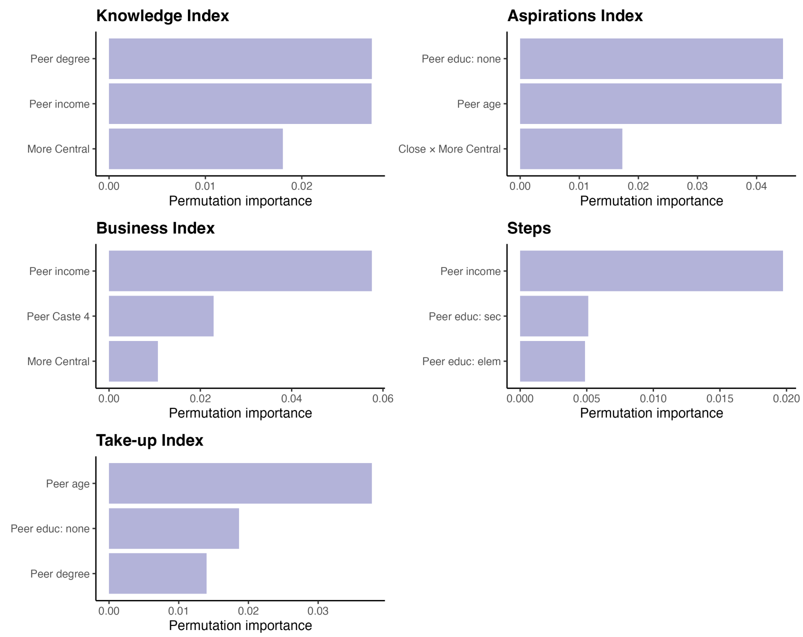
VARIABLES	(1) Knowledge	(2) Aspirations	(3) Intention	(4) Steps	(5) Take-up
Trained Alone (T1)	0.104	0.109	0.097	0.117	0.215**
	(0.094)	(0.112)	(0.101)	(0.099)	(0.094)
Popular x Close	0.099	0.452*	-0.021	-0.021	0.160
	(0.132)	(0.257)	(0.153)	(0.151)	(0.136)
Not Popular x Close	0.393***	0.340	0.368**	0.216	0.300**
	(0.131)	(0.266)	(0.158)	(0.153)	(0.141)
Popular x Far	0.241**	-0.066	-0.056	0.003	0.307***
	(0.119)	(0.131)	(0.152)	(0.154)	(0.107)
Not Popular x Far	0.203**	-0.098	0.171	0.158	0.204
	(0.098)	(0.123)	(0.122)	(0.108)	(0.124)
Observations	700	700	697	678	684
R-squared	0.051	0.020	0.026	0.027	0.034
p: Popular x Close = Not Popular x Close	0.0562	0.754	0.0483	0.208	0.402
p: Popular= Not Popular	0.173	0.703	0.0149	0.103	0.857
p: Close = Far	0.805	0.0120	0.356	0.887	0.806

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

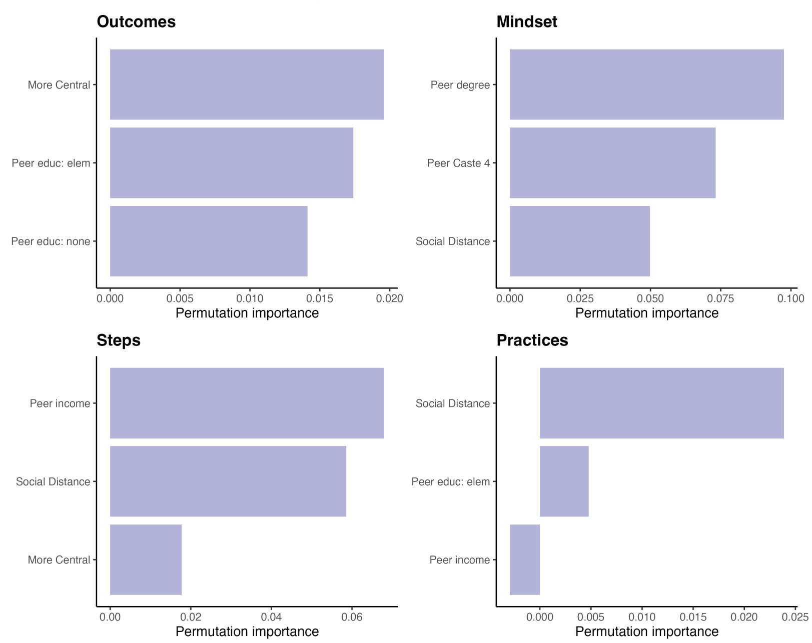
Notes: Each panel uses a different centrality measure and decomposes the paired treatment arms into four mutually exclusive cells based on whether the peer is popular (centrality $\geq$ 75th percentile in the village) and socially close. Coefficients are relative to within-village controls. All regressions control for the individual's own centrality. Robust standard errors in parentheses.

Figure SA.4: Outcome predictors using random forests

(A) Endline

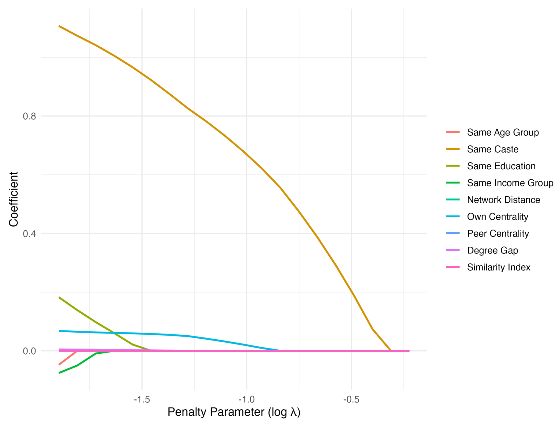


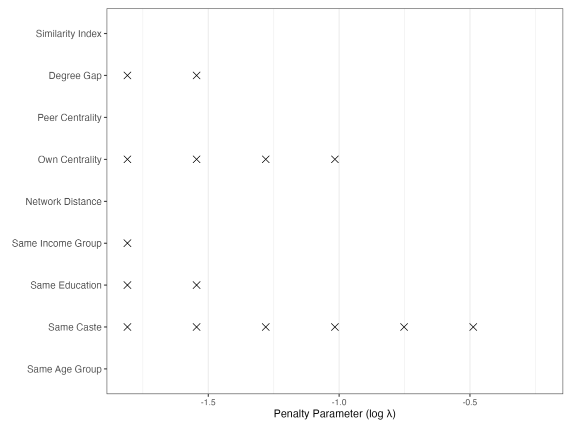
(B) Follow-up



Notes: The bars show the top three predictors of each outcome at endline (top) and follow-up (bottom), estimated using random forest models. Variable importance is measured by the increase in root mean squared error (RMSE) when the predictor is permuted. Outcomes are indices pre-specified in the analysis plan.

Figure SA.5: Lasso predictors of number of pooled network links





Notes: The left panel shows LASSO coefficient paths for predicting the number of links pooled within each pair who attended Treatment 3, regardless of assignment; the right panel shows when each variable enters the model as the penalty parameter (λ) decreases.

D Additional Follow-up Tables

Table SA.13: Effects on long-term outcomes by type of link (all)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.288	0.591***	0.288	0.148
	(0.193)	(0.196)	(0.201)	(0.149)
Trained Alone	0.191	0.453***	0.164	-0.0235
	(0.135)	(0.150)	(0.181)	(0.146)
Close x More Central	0.416	-0.0333	0.00358	0.278
	(0.517)	(0.191)	(0.405)	(0.243)
Close x Less Central	0.369	0.289	0.248	0.114
	(0.227)	(0.297)	(0.258)	(0.175)
Far x Central	-0.127	-0.175	-0.0268	-0.0904
	(0.165)	(0.147)	(0.193)	(0.161)
Far x Non Central	0.0438	0.348*	-0.0279	-0.00441
	(0.148)	(0.200)	(0.202)	(0.132)
Degree	0.00667	0.00674	0.00821	0.00367
	(0.0218)	(0.0256)	(0.0182)	(0.0251)
Constant	-0.0405	-0.0381	-0.0465	-0.0268
	(0.136)	(0.148)	(0.185)	(0.162)
Observations	566	570	570	570
R-squared	0.016	0.039	0.014	0.007
p: Central and Close vs Less Central and Close	0.942	0.324	0.552	0.508
p: Central and Close vs Central and Distant	0.347	0.471	0.942	0.103
p: Central and Close vs Less Central and Distant	0.500	0.122	0.943	0.272
p: Socially Close vs Distant	0.115	0.811	0.565	0.165
p: Less Central vs More Central	0.853	0.0463	0.606	0.803

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

Notes: This regression treats individuals in the pure control villages as the base category and includes an indicator for the spillover group and those who were intended to be paired but remained unmatched. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality. Standard errors are robust and clustered at the village level.

E Mechanism Tables

Table SA.14: Learning by type of pair

VARIABLES	(1) Knowledge Index	(2) Profit (Game)	(3) Profit (Business Plan)
Trained alone	0.0861		
	(0.0936)		
Close x More Central	0.281*	46.14	1.712e+06
	(0.146)	(81.27)	(1.125e+06)
Other Pair Types	0.177**	54.30	21,784
	(0.0874)	(43.04)	(247,688)
Degree	0.0459***	12.84	55,043
	(0.0122)	(8.028)	(53,121)
Constant	-0.220**	205.2***	238,772
	(0.0999)	(53.61)	(351,553)
Observations	700	427	427
R-squared	0.061	0.011	0.028
p: Close+Central vs Other Pairs	0.440	0.917	0.134
p: Close+Central vs T1	0.158		

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

Notes: The regression in Column 1 treats within-village controls as the base category while regressions in Columns 2 and 3 treat Treatment 1 as the base category since the games were only conducted for trainees. Standard errors are robust and clustered at the village level.

Table SA.15: Heterogeneous effects by own risk aversion

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	-0.105	0.325*	-0.0562	-0.106	0.0414
	(0.125)	(0.191)	(0.142)	(0.126)	(0.139)
Trained with Peer	-0.0680	0.319**	-0.146	-0.224*	0.0692
	(0.106)	(0.161)	(0.127)	(0.118)	(0.115)
Risk Averse (Binary)	-0.510***	0.129	-0.520***	-0.507***	-0.478***
	(0.148)	(0.152)	(0.147)	(0.148)	(0.144)
T1 x Risk Averse	0.398**	-0.272	0.309	0.399**	0.379*
	(0.191)	(0.242)	(0.206)	(0.198)	(0.198)
T2/T3 x Risk Averse	0.524***	-0.358*	0.449**	0.584***	0.313*
	(0.171)	(0.207)	(0.186)	(0.183)	(0.180)
Constant	0.272***	-0.102	0.278***	0.274***	0.254***
	(0.0809)	(0.110)	(0.0924)	(0.0805)	(0.0839)
Observations	664	664	661	642	647
R-squared	0.065	0.010	0.031	0.035	0.048

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

Notes: This regression treats within-village controls as the base category. Risk Aversion is measured on a scale of 1–6 using a series of choices between lotteries and fixed outcomes presented to participants in the baseline survey. Risk Averse (Binary) equals 1 if the individual's own risk aversion score is 5 or 6 (most risk-averse choices), and 0 otherwise. Robust standard errors in parentheses.

Table SA.16: Heterogeneous effects by peer risk aversion

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.0490	0.119	0.0904	0.102	0.217**
	(0.0829)	(0.107)	(0.0934)	(0.0883)	(0.0859)
T2 x Peer Less Risk Averse	-0.00392	-0.161	-0.0604	0.213	0.204
	(0.130)	(0.155)	(0.231)	(0.174)	(0.192)
T2 x Peer Same/More Risk Averse	0.142	0.204	0.0729	-0.0555	0.254**
	(0.0988)	(0.160)	(0.120)	(0.117)	(0.115)
T3 x Peer Less Risk Averse	0.0692	-0.292**	0.198	0.150	0.226
	(0.202)	(0.120)	(0.223)	(0.179)	(0.176)
T3 x Peer Same/More Risk Averse	0.196**	0.193	0.133	0.116	0.251**
	(0.0928)	(0.157)	(0.124)	(0.120)	(0.118)
Constant	0.0645	0.00642	0.0179	0.0243	0.0257
	(0.0600)	(0.0662)	(0.0630)	(0.0617)	(0.0607)
Observations	710	710	707	688	693
R-squared	0.040	0.009	0.013	0.022	0.028

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

Notes: This regression treats within-village controls as the base category. Peer Less Risk Averse equals 1 if the matched peer's risk aversion score is lower than the individual's own score. Peer Same/More Risk Averse equals 1 if the peer's score is the same or higher. Robust standard errors in parentheses.

Table SA.17: Treatment effects on network-based measures of saving

VARIABLES	(1) Number of Savings Groups	(2) Joined Cooperative in last year
Spillover	0.0310	0.0786
	(0.312)	(0.0533)
Trained Alone	0.144	0.112**
	(0.308)	(0.0478)
Trained with Peer	0.136	0.00104
	(0.298)	(0.0385)
Constant	1.864***	0.0929***
	(0.283)	(0.0329)
Observations	579	580
R-squared	0.002	0.020

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

Notes: This regression treats the pure control group as the base category. Standard errors are robust and clustered at the village level.

Table SA.18: Determinants of reporting encouragement as a benefit of pairing

VARIABLES	(1) Encouragement	(2) Encouragement
Peer caste = 2	-0.161	-0.151
	(0.128)	(0.123)
Peer caste = 3	-0.00793	-5.77e-05
	(0.125)	(0.120)
Peer caste = 4	-0.300**	-0.274**
	(0.135)	(0.130)
Peer education = 1	-0.239***	-0.240***
	(0.0872)	(0.0878)
Peer education = 2	-0.109	-0.112
	(0.0850)	(0.0854)
Peer education = 3	-0.147	-0.135
	(0.121)	(0.121)
Peer income	-8.61e-07	-8.96e-07
	(8.25e-07)	(8.04e-07)
Peer marital status = 1	0.0188	0.0305
	(0.131)	(0.126)
Peer marital status = 3	0.333	0.285
	(0.288)	(0.279)
Peer age	0.00241	0.00254
	(0.00402)	(0.00401)
Own network degree	-1.08e-05	
	(0.0130)	
Peer is more central	0.134*	
	(0.0707)	
Peer network degree		0.0302**
		(0.0122)
Constant	0.411	0.282
	(0.256)	(0.250)
Observations	247	247
R-squared	0.112	0.117

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

Notes: The dependent variable is an indicator equal to 1 if the respondent reported encouragement as a reason why pairing was beneficial. The sample is restricted to individuals assigned to the paired treatment arm and were actually trained with a peer. Robust standard errors in parentheses.

F Comparison with Pure Control

We replicate the main ITT estimates and peer heterogeneity specifications using pure control villages as the comparison group. Table SA.19 reports ITT effects with pure control as the base category. Tables SA.20–SA.22 confirm that patterns of peer heterogeneity are robust to the choice of control group.

Table SA.19: Impact of the training on immediate outcomes (pure control)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.687***	-0.165	0.232**	0.0970	0.186*
	(0.204)	(0.169)	(0.105)	(0.0811)	(0.106)
Treatment with Peer	0.595***	-0.182	0.0922	-0.104	0.0361
	(0.210)	(0.138)	(0.120)	(0.101)	(0.112)
Treatment with Peer + Connections Module	0.695***	-0.242	0.197*	0.0469	0.0925
	(0.201)	(0.154)	(0.101)	(0.0903)	(0.110)
Constant	-0.574***	0.290**	-0.124*	0.0291	0.0567
	(0.193)	(0.119)	(0.0722)	(0.0575)	(0.0789)
Observations	768	768	765	746	751
R-squared	0.127	0.006	0.010	0.006	0.006
p: Trained Alone vs with Peer	0.0923	0.895	0.256	0.0603	0.156
p: Trained Alone vs with Peer + CM	0.907	0.467	0.728	0.538	0.308
p: Treatment with Peer vs with Peer + CM	0.228	0.532	0.289	0.161	0.555

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

Notes: This regression treats individuals in pure control villages as the base category. Standard errors are robust and clustered at the village level.

Table SA.20: Heterogeneity by social distance and peer centrality (pure control)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.687***	-0.165	0.232**	0.0970	0.186*
	(0.204)	(0.169)	(0.105)	(0.0812)	(0.106)
Trained with Peer	0.994***	0.405	0.300	0.00921	-0.215
	(0.242)	(0.250)	(0.222)	(0.142)	(0.259)
Trained with Peer × Peer distance	-0.0282	-0.135***	-0.0210	0.00379	0.0471*
	(0.0231)	(0.0327)	(0.0403)	(0.0276)	(0.0275)
Trained with Peer × Peer degree	-0.0190	-0.0295	0.00125	0.0105	0.0480*
	(0.0179)	(0.0253)	(0.0234)	(0.0139)	(0.0241)
Constant	-0.574***	0.290**	-0.124*	0.0291	0.0567
	(0.193)	(0.119)	(0.0723)	(0.0576)	(0.0791)
Observations	753	753	750	731	737
R-squared	0.166	0.016	0.019	0.019	0.029

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

Notes: This regression treats individuals in pure control villages as the base category. Standard errors are robust and clustered at the village level.

Table SA.21: Heterogeneity by social closeness and relative centrality (pure control)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.673***	-0.184	0.210*	0.0732	0.141
	(0.204)	(0.168)	(0.103)	(0.0786)	(0.103)
Trained with Peer	0.803***	-0.434**	0.205	0.0412	0.112
	(0.205)	(0.165)	(0.121)	(0.0943)	(0.113)
Trained with Peer (Close)	-0.00431	0.475***	0.0355	-0.0375	-0.0464
	(0.0807)	(0.161)	(0.154)	(0.119)	(0.0997)
Trained with Peer (More Central)	-0.0425	0.157	-0.0134	0.0390	0.159*
	(0.0939)	(0.168)	(0.126)	(0.137)	(0.0846)
Degree	0.0180	0.0227	0.0352*	0.0293*	0.0308*
	(0.0127)	(0.0288)	(0.0174)	(0.0164)	(0.0161)
Constant	-0.659***	0.176	-0.291**	-0.104	-0.0821
	(0.199)	(0.139)	(0.122)	(0.108)	(0.127)
Observations	756	756	753	734	740
R-squared	0.167	0.020	0.026	0.023	0.028

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

Notes: This regression treats individuals in pure control villages as the base category. Standard errors are robust and clustered at the village level.

Table SA.22: Heterogeneity by peer type (pure control)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.672***	-0.188	0.209*	0.0741	0.144
	(0.204)	(0.168)	(0.103)	(0.0777)	(0.102)
Close x More Central	0.852***	0.460	0.357*	0.0531	0.308***
	(0.245)	(0.284)	(0.181)	(0.193)	(0.106)
Other Pair Types	0.772***	-0.329**	0.185*	0.0422	0.133
	(0.201)	(0.148)	(0.102)	(0.0800)	(0.112)
Degree	0.0199*	0.0308	0.0375**	0.0275*	0.0260*
	(0.0117)	(0.0264)	(0.0160)	(0.0154)	(0.0149)
Constant	-0.669***	0.137	-0.302**	-0.0951	-0.0589
	(0.198)	(0.134)	(0.116)	(0.103)	(0.121)
Observations	756	756	753	734	740
R-squared	0.167	0.024	0.027	0.023	0.027
p: Central and Close vs Other Pairs	0.607	0.00621	0.410	0.961	0.0632
p: Central and Close vs T1	0.260	0.0249	0.415	0.906	0.0707

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

Notes: This regression treats individuals in pure control villages as the base category. Standard errors are robust and clustered at the village level.

G PDS-Lasso Robustness Checks

SA.1 Short-Term Outcomes

Table SA.23: Effects on endline outcomes by peer type (PDS-Lasso)

(a) By Social Distance and Peer Centrality

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.112	0.129	0.121	0.115	0.193**
	(0.0913)	(0.0941)	(0.0973)	(0.0940)	(0.0906)
Trained with Peer	0.250	0.399*	0.00269	-0.0680	-0.290
	(0.164)	(0.216)	(0.205)	(0.181)	(0.211)
Trained with Peer × Peer distance	-0.00326	-0.0565*	0.00334	0.0182	0.0629**
	(0.0244)	(0.0334)	(0.0317)	(0.0277)	(0.0275)
Trained with Peer × Peer degree	-0.00332	-0.0377*	0.0145	0.0166	0.0580***
	(0.0187)	(0.0223)	(0.0220)	(0.0196)	(0.0213)
Constant	0.362**	-0.491***	0.890***	0.834***	1.032***
	(0.169)	(0.0878)	(0.190)	(0.169)	(0.144)
Observations	678	678	675	656	662
Number of groups	0	0	0	0	0

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

(b) By Social Closeness and More/Less Central

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.105	0.119	0.0990	0.0992	0.183**
	(0.0910)	(0.0946)	(0.0966)	(0.0947)	(0.0901)
Trained with Peer	0.214**	-0.0201	0.0474	0.0723	0.174
	(0.0912)	(0.116)	(0.108)	(0.106)	(0.109)
Trained with Peer (Close)	-0.0657	0.176	-0.0472	-0.0820	-0.107
	(0.0938)	(0.133)	(0.120)	(0.114)	(0.108)
Trained with Peer (More Central)	0.0677	-0.0570	0.0680	0.0307	0.193*
	(0.0944)	(0.127)	(0.121)	(0.120)	(0.109)
Constant	0.197	-0.430***	0.542**	0.728***	0.863***
	(0.193)	(0.112)	(0.231)	(0.197)	(0.164)
Observations	681	681	678	659	665
Number of groups	0	0	0	0	0

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

(c) By Social Closeness Interacted with More/Less Central

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.106	0.117	0.119	0.109	0.198**
	(0.0910)	(0.0940)	(0.0976)	(0.0944)	(0.0910)
Close x More Central	0.284**	0.219	0.151	0.0125	0.334***
	(0.143)	(0.199)	(0.146)	(0.158)	(0.0972)
Other Pair Types	0.201**	-0.0320	0.0763	0.0762	0.216**
	(0.0849)	(0.0920)	(0.0929)	(0.0907)	(0.0940)
Constant	0.383**	-0.469***	0.895***	0.848***	0.984***
	(0.167)	(0.0881)	(0.195)	(0.169)	(0.142)
Observations	681	681	678	659	665
Number of groups	0	0	0	0	0
p: Central and Close vs Other Pairs	0.530	0.199	0.602	0.678	0.182
p: Central and Close vs T1	0.193	0.607	0.830	0.535	0.112

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

Notes: These regressions use post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be correlated with treatment status. Covariates include income, source of income, age, education, caste, and number of network connections. Within-village controls are the base category. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality.

SA.2 Long-Term Outcomes

The following tables replicate the main follow-up results (Table 3 in the paper) using post double selection Lasso (Belloni et al. 2014) to control for baseline covariates.

Table SA.24: Effects on long-term outcomes by peer type (PDS-Lasso)

(a) By Social Distance and Peer Centrality

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained alone	0.148	0.472***	0.163	0.00473
	(0.132)	(0.146)	(0.179)	(0.135)
Trained with Peer	0.495	0.272	0.279	0.339*
	(0.303)	(0.355)	(0.338)	(0.204)
Trained with Peer × Peer distance	-0.0995**	0.0288	-0.0565	-0.101**
	(0.0445)	(0.0692)	(0.0476)	(0.0439)
Trained with Peer × Peer degree	-0.0166	-0.0458	-0.0109	0.00996
	(0.0266)	(0.0289)	(0.0306)	(0.0218)
Constant	0.0281	0.729***	-0.355*	0.0157
	(0.181)	(0.179)	(0.188)	(0.0886)
Observations	561	565	565	565
Number of groups	0	0	0	0

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

(b) By Social Closeness and More/Less Central

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained alone	0.154	0.476***	0.161	0.00459
	(0.125)	(0.146)	(0.174)	(0.136)
Trained with Peer	-0.0549	0.330*	0.0260	-0.0195
	(0.170)	(0.196)	(0.200)	(0.126)
Trained with Peer (Close)	0.401*	-0.0194	0.0795	0.210
	(0.212)	(0.180)	(0.197)	(0.169)
Trained with Peer (More Central)	-0.0849	-0.435**	-0.0814	-0.00728
	(0.227)	(0.179)	(0.202)	(0.129)
Constant	-0.0470	0.721***	-0.297	-0.0323
	(0.257)	(0.200)	(0.192)	(0.158)
Observations	565	569	569	569
Number of groups	0	0	0	0

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

(c) By Social Closeness Interacted with More/Less Central

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained alone	0.163	0.470***	0.164	0.00747
	(0.129)	(0.146)	(0.177)	(0.137)
Close x Less Central	0.317	0.257	0.223	0.137
	(0.200)	(0.275)	(0.236)	(0.166)
Other Pair Types	-0.0104	0.121	-0.0232	0.0307
	(0.124)	(0.137)	(0.172)	(0.128)
Constant	-0.107	0.682***	-0.302	-0.0479
	(0.239)	(0.191)	(0.201)	(0.154)
Observations	565	569	569	569
Number of groups	0	0	0	0
p: Less Central and Close vs Other Pairs	0.147	0.590	0.204	0.541
p: Less Central and Close vs T1	0.440	0.439	0.772	0.404

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

Notes: These regressions use post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be correlated with treatment status. Covariates include income, source of income, age, education, caste, and number of network connections. Individuals in the pure control villages are the base category. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality.

Table SA.25: Effects on long-term outcomes (PDS-Lasso)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.198	0.608***	0.275	0.159
	(0.187)	(0.189)	(0.188)	(0.147)
Trained Alone	0.167	0.468***	0.173	0.00199
	(0.132)	(0.149)	(0.168)	(0.143)
Trained with Peer	0.0726	0.209*	0.0217	-0.000431
	(0.109)	(0.122)	(0.156)	(0.117)
Constant	-0.107	0.709***	0.402	-0.0210
	(0.189)	(0.161)	(0.278)	(0.0899)
Observations	573	577	577	577
Number of groups	0	0	0	0
p: Trained Alone v/s Spillover	0.863	0.458	0.413	0.279
p: Trained Alone v/s Trained with Peer	0.388	0.0746	0.0625	0.982
p: Trained with Peer v/s Spillover	0.433	0.0221	0.0345	0.191

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

Notes: This regression treats individuals in the pure control group as the base category and includes an indicator for the spillover group. Standard errors are robust and clustered at the village level. We implement PDS Lasso (Belloni et al. 2014) including baseline controls for income (level and sources), age, education (elementary, secondary, higher, university, informal), and caste (categories 1–6).

H Local Average Treatment Effects

Table SA.26: Impact of the training on immediate outcomes within treated villages (LATE)

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained Alone	0.474	0.538	0.328	0.313	0.981**
	(0.392)	(0.481)	(0.444)	(0.452)	(0.431)
Trained with Peer	0.382	0.445	0.166	0.0831	0.737*
	(0.380)	(0.471)	(0.434)	(0.443)	(0.425)
Constant	-0.256	-0.295	-0.133	-0.0930	-0.508
	(0.300)	(0.356)	(0.334)	(0.343)	(0.327)
Observations	710	710	707	688	693
R-squared	0.160	-0.017	0.039	0.029	0.099
Under-ID p(KP LM)	2.87e-06	2.87e-06	3.23e-06	5.49e-06	6.89e-06
First-stage F: Trained Alone	127.5	127.5	124.1	114.1	115
First-stage F: Trained with peer	166.7	166.7	163.5	149.7	150.6
p: Trained Alone = Trained with peer	0.397	0.599	0.263	0.0891	0.0587
q: Trained Alone	0.542	0.542	0.642	0.642	0.130
q: Trained with Peer	0.849	0.849	1	1	0.713

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

Notes: This regression treats the control sample within treated villages as the base category. Local average treatment effects (LATE) are estimated by instrumenting actual treatment status with assigned treatment. Robust standard errors in parentheses. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes.

Table SA.27: Impact of the training on follow-up outcomes (LATE)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	-0.0432	0.292	-0.0710	0.0555
	(0.221)	(0.231)	(0.264)	(0.229)
Trained Alone	0.244	0.552**	0.229	-0.0620
	(0.198)	(0.229)	(0.219)	(0.213)
Trained with Peer	0.225*	0.328**	0.148	0.0437
	(0.129)	(0.156)	(0.169)	(0.119)
Constant	0	-0	0	0
	(0.0685)	(0.0706)	(0.154)	(0.0840)
Observations	576	580	580	580
R-squared	0.004	0.014	0.007	0.001
Under-ID p(KP LM)	0.00103	0.000730	0.000730	0.000730
First-stage F: Trained Alone	95.09	100.8	100.8	100.8
First-stage F: Trained with peer	150.6	150.6	150.6	150.6
p: Trained Alone = Trained with peer	0.924	0.377	0.633	0.567
q: Spillover	1	1	1	1
q: Trained Alone	1	1	1	1
q: Trained with Peer	1	1	1	1

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

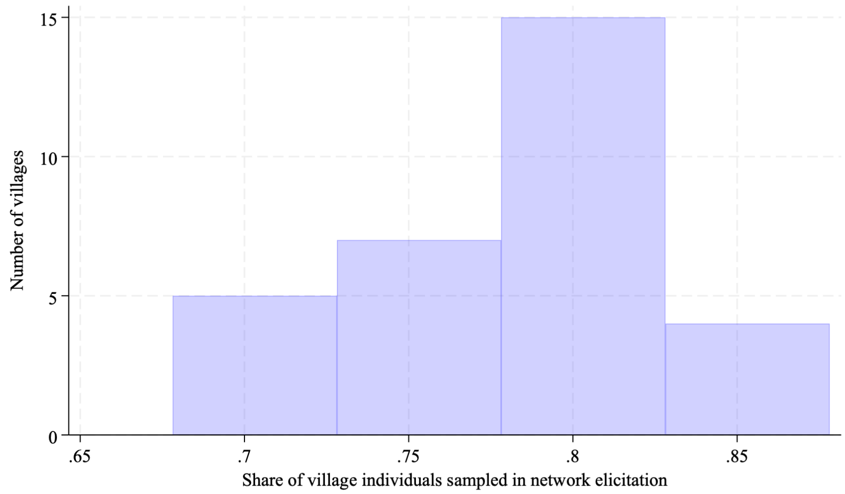
Notes: This regression treats the pure control group as the base category and excludes within village controls. Standard errors clustered at village level. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes.

I Robustness to Network Sampling

Figure SA.6 shows the proportion of village populations for which we elicited network data. On average, this rate is 78%. We correct our measure of degree centrality to account for unobserved links by dividing the observed degree by the village-specific sampling proportion. The intuition is as follows: if individual i has d_i links in an observed network with x nodes, then in a village with population y they are predicted to have $\frac{d_i}{x} \times y$ links. Figure SA.7 shows the distribution of village densities based on this corrected measure. Density is higher than the earlier conservative estimate but remains low overall, indicating sparse networks.

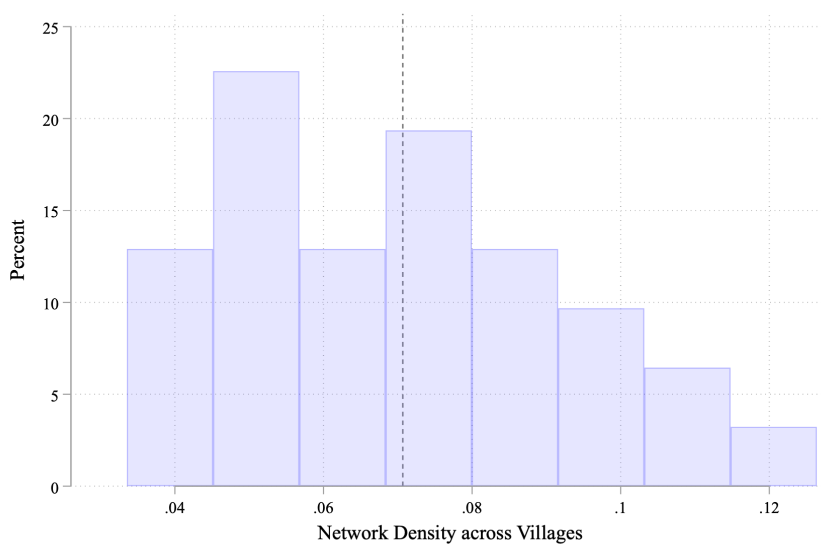
Next, we present endline and follow-up results using this corrected measure of centrality in Table SA.28 and Table SA.29. The results show that the positive effects of close and central peers on aspirations at endline, and of close and less central peers at follow-up, are robust to this correction.

Figure SA.6: Fraction of individuals per village sampled in network elicitation



Notes: The figure plots the distribution of the share of each village's population for which network data were elicited. Village population is defined as the union of all individuals who had network data collected or were listed as contacts by others but do not have network data. The sampling rate for each village is the ratio of individuals with network data to this constructed village population.

Figure SA.7: Network density across villages



Notes: The figure shows the distribution of village-level network density computed using the degree measure corrected for sampling coverage. For each village, we first adjust individual degrees accounting for individuals who are not included in the network elicitation, and then calculate density based on these corrected degrees.

Table SA.28: Effects on endline outcomes by peer type

(a) By Social Distance and Peer Centrality

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.142	0.0895	0.120	0.131	0.201**
	(0.112)	(0.0785)	(0.112)	(0.108)	(0.0802)
Trained with Peer	0.548***	0.437**	0.187	0.0616	-0.0757
	(0.194)	(0.178)	(0.240)	(0.205)	(0.166)
Trained with Peer × Peer distance	-0.0304	-0.0811***	-0.0227	0.00355	0.0220
	(0.0279)	(0.0289)	(0.0379)	(0.0317)	(0.0206)
Trained with Peer × Peer degree	-0.0227	-0.0147	0.00164	0.00554	0.0276**
	(0.0174)	(0.0146)	(0.0213)	(0.0173)	(0.0139)
Constant	0.647***	-0.229***	0.133	-0.0294	-0.0961*
	(0.0881)	(0.0519)	(0.0830)	(0.0827)	(0.0568)
Observations	695	695	692	673	679
R-squared	0.052	0.012	0.013	0.020	0.034

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

(b) By Social Closeness and More/Less Central

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.112	0.0802	0.0858	0.116	0.162**
	(0.111)	(0.0793)	(0.111)	(0.108)	(0.0763)
Trained with Peer	0.270**	-0.0563	0.0656	0.0877	0.104
	(0.108)	(0.0936)	(0.131)	(0.125)	(0.0883)
Trained with Peer (Close)	-0.0440	0.297**	0.0124	-0.0388	-0.0135
	(0.111)	(0.123)	(0.141)	(0.131)	(0.0785)
Trained with Peer (More Central)	0.00832	0.0606	0.0306	0.0310	0.141*
	(0.108)	(0.114)	(0.143)	(0.134)	(0.0850)
Degree (Alternative)	0.0374***	0.000820	0.0466***	0.0203	0.0248***
	(0.0115)	(0.0112)	(0.0136)	(0.0136)	(0.00860)
Constant	0.419***	-0.234***	-0.151	-0.154	-0.248***
	(0.118)	(0.0809)	(0.124)	(0.121)	(0.0839)
Observations	700	700	697	678	684
R-squared	0.062	0.014	0.028	0.022	0.037

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

(c) By Social Closeness Interacted with More/Less Central

VARIABLES	(1) Knowledge Index	(2) Aspirations Index	(3) Business Index	(4) Additional Steps	(5) Take-up Index
Trained alone	0.112	0.0768	0.0854	0.116	0.164**
	(0.111)	(0.0792)	(0.111)	(0.108)	(0.0762)
Close x More Central	0.353**	0.483**	0.255	0.0909	0.280***
	(0.163)	(0.218)	(0.170)	(0.180)	(0.0845)
Other Pair Types	0.241**	-0.00581	0.0501	0.0859	0.135*
	(0.101)	(0.0732)	(0.110)	(0.107)	(0.0732)
Degree (Alternative)	0.0372***	0.00609	0.0471***	0.0190	0.0215***
	(0.0109)	(0.0108)	(0.0128)	(0.0125)	(0.00796)
Constant	0.420***	-0.266***	-0.154	-0.146	-0.228***
	(0.115)	(0.0797)	(0.119)	(0.116)	(0.0807)
Observations	700	700	697	678	684
R-squared	0.063	0.019	0.029	0.022	0.036
p: Central and Close vs Other Pairs	0.450	0.0248	0.223	0.977	0.0671
p: Central and Close vs T1	0.120	0.0649	0.313	0.884	0.156

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

Notes: These regressions treat individuals in the pure control villages as the base category and include an indicator for those who were intended to be paired but remained unmatched. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality. Standard errors are robust and clustered at the village level.

Table SA.29: Effects on long-term outcomes

(a) By Social Distance and Peer Centrality

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.194	0.490***	0.173	-0.0201
	(0.137)	(0.164)	(0.188)	(0.146)
Trained with Peer	0.565*	0.427	0.511	0.326
	(0.324)	(0.387)	(0.342)	(0.222)
Trained with Peer × Peer distance	-0.0994**	0.00975	-0.102*	-0.101**
	(0.0466)	(0.0707)	(0.0524)	(0.0454)
Trained with Peer × Peer degree	-0.0163	-0.0489*	-0.0203	0.00598
	(0.0230)	(0.0239)	(0.0232)	(0.0183)
Constant	0	-0	-0	-0
	(0.0702)	(0.0723)	(0.158)	(0.0860)
Observations	564	568	568	568
R-squared	0.013	0.035	0.018	0.012

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

(b) By Social Closeness and More/Less Central

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.196	0.455***	0.165	-0.0224
	(0.135)	(0.149)	(0.180)	(0.145)
Trained with Peer	0.0245	0.326	0.00508	-0.0332
	(0.174)	(0.195)	(0.204)	(0.126)
Trained with Peer (Close)	0.424*	0.0252	0.182	0.221
	(0.218)	(0.183)	(0.250)	(0.173)
Trained with Peer (More Central)	-0.121	-0.467**	-0.0837	-0.0129
	(0.229)	(0.182)	(0.196)	(0.136)
Degree (Alternative)	-0.00164	0.00244	0.00337	0.000900
	(0.0178)	(0.0204)	(0.0157)	(0.0199)
Constant	0.00216	-0.0198	-0.0266	-0.0143
	(0.137)	(0.147)	(0.185)	(0.159)
Observations	566	570	570	570
R-squared	0.015	0.038	0.013	0.006

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

(c) By Social Closeness Interacted with More/Less Central

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.287	0.591***	0.286	0.147
	(0.192)	(0.197)	(0.201)	(0.148)
Trained Alone	0.193	0.447***	0.165	-0.0237
	(0.135)	(0.149)	(0.180)	(0.145)
Close x Less Central	0.380*	0.270	0.256	0.114
	(0.223)	(0.294)	(0.259)	(0.176)
Other Pair Types	0.0442	0.113	-0.0232	0.0100
	(0.135)	(0.141)	(0.168)	(0.126)
Degree (Alternative)	0.00116	0.0106	0.00324	0.00233
	(0.0174)	(0.0185)	(0.0165)	(0.0191)
Constant	-0.0148	-0.0692	-0.0258	-0.0229
	(0.135)	(0.136)	(0.187)	(0.155)
Observations	566	570	570	570
R-squared	0.011	0.031	0.014	0.004
p: Close x Less Central vs Other Pairs	0.182	0.555	0.203	0.562
p: Close x Less Central vs T1	0.393	0.542	0.708	0.393

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

Notes: This regression treats individuals in the pure control villages as the base category and includes an indicator for the spillover group and those who were intended to be paired but remained unmatched. We additionally control for the individual's own degree centrality in regressions that include relative comparisons between own and peer centrality. Standard errors are robust and clustered at the village level.

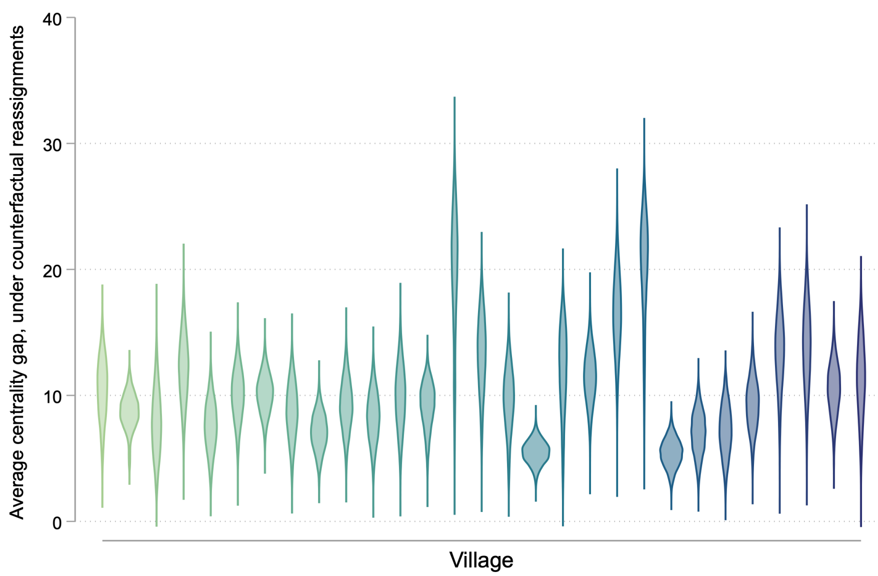
J Peer Effects and Feasibility of Peer Pairings

Table SA.30: Identification of peer effects (robustness)

VARIABLES	(1) Outcome (Follow-up)	(2) Outcome (Follow-up)	(3) Outcome (Follow-up)
Peer Outcome (Endline)	0.276*	0.343*	0.381*
	(0.163)	(0.183)	(0.210)
Peer Outcome (Follow-up)	0.141*	0.151*	-0.0829
	(0.0762)	(0.0776)	(0.0776)
Close X Peer Outcome	0.0238	-0.0467	-0.117
	(0.188)	(0.188)	(0.200)
Peer Degree X Peer Outcome	-0.0646**	-0.0581*	-0.0594*
	(0.0306)	(0.0321)	(0.0347)
Own Outcome (Endline)	0.295***	0.101	0.107
	(0.0757)	(0.0817)	(0.100)
Constant	-0.0843	0.736	0.539
	(0.0658)	(0.745)	(0.842)
Observations	186	186	186
R-squared	0.083	0.192	0.360

Notes: The table reports the effect of the matched peer's endline outcome on the individual's own follow-up outcome, additionally controlling for the peer's own outcome in the follow up. Close × Peer Outcome interacts the peer outcome with an indicator for whether the peer was socially close. Peer Degree × Peer Outcome interacts the peer outcome with the peer's degree centrality. Columns (2) and (3) control for the individual's and peer's age, income, education, caste, and network degree. Column (3) additionally includes village fixed effects.

Figure SA.8: Feasibility of counterfactual pair reassignments



Notes: The figure plots the distribution of the *average within-village degree-centrality gap* across 10,000 random reassignments of peer dyads (without replacement), holding the network structure fixed.

K Robustness to Attrition: IPW-Weighted Results

The following tables replicate the main follow-up results using Inverse Probability Weighting (IPW) to account for potential attrition bias between the endline and follow-up surveys. We estimate the probability of being observed at follow-up, conditional on being in the endline sample, using a probit model. The covariates include treatment assignment indicators and all baseline characteristics used in the balance tables. The IPW weight for each individual is the inverse of their predicted probability of being observed, $w_i = 1/\hat{p}_i$. All follow-up regressions are then re-estimated using these weights.

Table SA.31: Follow-up effects on main outcomes (IPW)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Spillover	0.273	0.564***	0.250	0.150
	(0.194)	(0.199)	(0.207)	(0.154)
Trained Alone	0.182	0.417**	0.127	0.00829
	(0.138)	(0.151)	(0.185)	(0.149)
Trained with Peer	0.127	0.241*	0.0143	0.00498
	(0.116)	(0.137)	(0.171)	(0.121)
Constant	-0.0115	-0.0235	-0.00805	-0.01000
	(0.0689)	(0.0726)	(0.158)	(0.0909)
Observations	561	565	565	565
R-squared	0.006	0.025	0.007	0.003
p: Spillover vs T1	0.632	0.503	0.409	0.376
p: T1 vs Paired	0.625	0.297	0.289	0.977
p: Spillover vs Paired	0.362	0.128	0.0890	0.285

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

Notes: Base category is pure control. Observations are weighted by inverse probability of survey response. Standard errors clustered at village level.

Table SA.32: Follow-up effects by social distance and peer centrality (IPW)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.182	0.417**	0.127	0.00829
	(0.139)	(0.151)	(0.185)	(0.150)
Trained with Peer	0.580*	0.363	0.514	0.354
	(0.326)	(0.369)	(0.359)	(0.219)
Trained with Peer × Peer distance	-0.0992*	0.0200	-0.104*	-0.101**
	(0.0483)	(0.0705)	(0.0537)	(0.0456)
Trained with Peer × Peer degree	-0.0205	-0.0561	-0.0256	0.00614
	(0.0267)	(0.0338)	(0.0305)	(0.0237)
Constant	-0.0115	-0.0235	-0.00805	-0.01000
	(0.0691)	(0.0728)	(0.159)	(0.0912)
Observations	552	556	556	556
R-squared	0.012	0.033	0.015	0.012

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

Notes: Base category is pure control. Observations are weighted by inverse probability of survey response. Standard errors clustered at village level.

Table SA.33: Follow-up effects by social closeness and More/Less Central (IPW)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.182	0.410**	0.121	0.00980
	(0.138)	(0.148)	(0.183)	(0.149)
Trained with Peer	0.0558	0.344	-0.000333	-0.0319
	(0.170)	(0.212)	(0.211)	(0.132)
Trained with Peer (Close)	0.414*	0.0113	0.186	0.228
	(0.208)	(0.180)	(0.252)	(0.176)
Trained with Peer (More Central)	-0.142	-0.467**	-0.0903	0.00702
	(0.218)	(0.179)	(0.204)	(0.136)
Degree	-0.000434	0.0115	0.0103	-0.00247
	(0.0222)	(0.0250)	(0.0182)	(0.0256)
Constant	-0.00938	-0.0792	-0.0577	0.00196
	(0.131)	(0.143)	(0.188)	(0.165)
Observations	554	558	558	558
R-squared	0.015	0.038	0.010	0.006

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

Notes: Base category is pure control. Observations are weighted by inverse probability of survey response. Standard errors clustered at village level.

Table SA.34: Follow-up effects by friend type (IPW)

VARIABLES	(1) Outcomes Index	(2) Steps Index	(3) Mindset Index	(4) Business Practices
Trained Alone	0.180	0.404**	0.121	0.00929
	(0.137)	(0.148)	(0.183)	(0.148)
Close x Less Central	0.403*	0.266	0.275	0.145
	(0.230)	(0.287)	(0.271)	(0.180)
Other Pair Types	0.0625	0.123	-0.0356	0.0186
	(0.135)	(0.151)	(0.172)	(0.133)
Degree	0.00345	0.0221	0.00991	-0.00163
	(0.0220)	(0.0228)	(0.0194)	(0.0245)
Constant	-0.0282	-0.131	-0.0560	-0.00211
	(0.130)	(0.132)	(0.191)	(0.161)
Observations	554	558	558	558
R-squared	0.010	0.030	0.011	0.004
p: Less Central Friend vs Other Pairs	0.202	0.582	0.178	0.492
p: Less Central Friend vs T1	0.326	0.636	0.552	0.419

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

Notes: Base category is pure control. Observations are weighted by inverse probability of survey response. Standard errors clustered at village level.

L Endline Index Components

SA.1 Outcome Descriptions

Knowledge Index. This index measures comprehension of business concepts covered during training. It comprises five quiz questions, each coded as 1 if answered correctly and 0 otherwise:

1. "What do you understand by a business?" (Correct: intended to make a profit)
2. "What characteristics are required to be a successful entrepreneur?" (Correct: all listed characteristics)
3. "What do you mean by fixed assets?" (Correct: lasting more than 1 year)
4. "What sector does a beauty parlour come under?" (Correct: service industry)
5. "Above what break-even percentage does the business become risky?" (Correct: 60%)

Aspirations Index. This index captures business-related aspirations following the elicitation procedure in Bernard & Seyoum Taffesse (2014). Respondents were asked: "What is the level you would like to reach?" for each of the following:

1. Yearly agricultural investment (NPR)
2. Yearly non-agricultural business investment (NPR)
3. Monthly income (NPR)
4. Monthly savings (NPR)

All monetary variables are winsorized at the 1st and 99th percentiles.

Business Intention Index. This index captures forward-looking entrepreneurial intentions:

1. Ready to invest in a business (ordinal: 0=not ready, 1=maybe, 2=of course)
2. Would submit business plan for a competition (binary)
3. Likelihood of starting a business (scale 1–5)

Steps Index. This index captures concrete preparatory actions for starting a business:

1. Intends to open a savings account for business purposes (binary)
2. Has taken a loan for business purposes (binary)

Take-up Index. This index measures willingness to seek additional support post-training:

1. Would attend additional training (binary)
2. Would attend mentoring workshops (binary)
3. Would seek advice from community members (binary)
4. Willingness to pay for trainer services (NPR, winsorized at 99th percentile)
5. Willingness to pay for mentoring services (NPR, winsorized at 99th percentile)

Table SA.35: Endline knowledge index components: ITT effects

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4	(5) Q5
Trained alone	0.0228	-0.00387	0.0228	0.0843*	-0.00182
	(0.0310)	(0.0525)	(0.0513)	(0.0470)	(0.0424)
Treatment with Peer	0.0367	-0.0773	-0.0425	0.0393	0.0215
	(0.0301)	(0.0523)	(0.0526)	(0.0487)	(0.0416)
Treatment with Peer + Connections Module	0.0457	-0.0301	0.0505	0.0705	0.0325
	(0.0302)	(0.0545)	(0.0532)	(0.0493)	(0.0426)
Constant	0.897***	0.477***	0.640***	0.704***	0.807***
	(0.0237)	(0.0378)	(0.0376)	(0.0360)	(0.0307)
Observations	690	710	679	680	688
R-squared	0.004	0.004	0.005	0.006	0.001
p: T1 vs T2	0.609	0.153	0.199	0.313	0.565
p: T1 vs T3	0.401	0.625	0.589	0.760	0.408
p: T2 vs T3	0.733	0.377	0.0773	0.507	0.787
q: T1	1	1	1	0.577	1
q: T2	1	1	1	1	1
q: T3	0.621	0.621	0.621	0.621	0.621

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

Notes: Base category is within-village controls. Robust standard errors in parentheses. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.36: Endline knowledge index components: effects by peer type

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4	(5) Q5
Trained alone	0.0181	-0.0110	0.0110	0.0763	-0.00498
	(0.0310)	(0.0527)	(0.0513)	(0.0468)	(0.0426)
Close x More Central	0.0255	-0.00289	0.154**	0.221***	0.115**
	(0.0497)	(0.0900)	(0.0759)	(0.0563)	(0.0536)
Other Pair Types	0.0527*	-0.0579	0.0675	0.117***	0.0802**
	(0.0272)	(0.0513)	(0.0490)	(0.0442)	(0.0376)
Degree	0.0100***	0.00879	0.0239***	0.0170***	0.00492
	(0.00363)	(0.00821)	(0.00730)	(0.00643)	(0.00582)
Constant	0.848***	0.435***	0.523***	0.620***	0.783***
	(0.0321)	(0.0545)	(0.0527)	(0.0499)	(0.0426)
Observations	680	700	669	670	678
R-squared	0.017	0.004	0.067	0.073	0.040
p: Close+Central vs Other	0.553	0.536	0.241	0.0423	0.483
p: Close+Central vs T1	0.878	0.928	0.0572	0.00684	0.0245
q: Close+Central	0.437	0.639	0.0620	0.00100	0.0620
q: Other Pairs	0.0770	0.117	0.0980	0.0440	0.0720

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

Notes: Base category is within-village controls. “Close x Central” indicates pairing with a socially close peer who is more central. Robust standard errors in parentheses.

Table SA.37: Endline aspirations index components: ITT effects

VARIABLES	(1) Agricultural Aspirations	(2) Non-Agricultural Aspirations	(3) Income Aspirations	(4) Savings Aspirations
Trained alone	20,470	2,054	5,185	1,759
	(22,327)	(31,803)	(7,378)	(1,417)
Treatment with Peer	-19,451	-11,492	3,646	2,092
	(20,807)	(26,730)	(7,965)	(1,453)
Treatment with Peer + Connections Module	3,494	62,137	-4,437	993.1
	(22,719)	(40,257)	(7,394)	(1,255)
Constant	225,369***	103,226***	57,768***	5,883***
	(14,846)	(20,879)	(4,998)	(844.5)
Observations	707	604	709	706
R-squared	0.005	0.009	0.002	0.003
p: T1 vs T2	0.0719	0.643	0.852	0.839
p: T1 vs T3	0.479	0.153	0.211	0.602
p: T2 vs T3	0.309	0.0547	0.328	0.465
q: T1	1	1	1	1
q: T2	1	1	1	1
q: T3	1	0.973	1	1

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

Notes: Base category is within-village controls. Robust standard errors in parentheses. Continuous variables are winsorized at the 1st and 99th percentiles. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.38: Endline aspirations index components: effects by peer type

VARIABLES	(1) Agricultural Aspirations	(2) Non-Agricultural Aspirations	(3) Income Aspirations	(4) Savings Aspirations
Trained alone	16,350	935.4	3,505	1,790
	(22,406)	(32,113)	(7,528)	(1,405)
Close x More Central	29,687	144,836*	52,121**	2,016
	(46,016)	(85,322)	(24,905)	(2,743)
Other Pair Types	-12,886	-13,560	-7,263	1,691
	(21,125)	(26,926)	(6,636)	(1,352)
Degree	5,443	-6,186	2,030*	-94.04
	(3,367)	(4,961)	(1,231)	(285.0)
Constant	199,299***	132,438***	48,003***	6,335***
	(21,483)	(29,303)	(7,466)	(1,586)
Observations	697	597	699	696
R-squared	0.008	0.021	0.035	0.003
p: Close+Central vs Other	0.356	0.0614	0.0165	0.907
p: Close+Central vs T1	0.775	0.0959	0.0523	0.936
q: Close+Central	0.351	0.173	0.173	0.351
q: Other Pairs	1	1	1	1

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

Notes: Base category is within-village controls. "Close x Central" indicates pairing with a socially close peer who is more central. Robust standard errors in parentheses.

Table SA.39: Endline business intention index components: ITT effects

VARIABLES	(1) Ready to Invest	(2) Submit Plan	(3) Business Likelihood
Trained alone	0.0886	0.0452	0.0959
	(0.0841)	(0.0476)	(0.134)
Treatment with Peer	-0.0335	0.0131	-0.111
	(0.0876)	(0.0485)	(0.136)
Treatment with Peer + Connections Module	0.00500	0.0571	0.246*
	(0.0888)	(0.0490)	(0.137)
Constant	1.320***	0.708***	3.586***
	(0.0607)	(0.0349)	(0.0957)
Observations	706	681	564
R-squared	0.003	0.003	0.012
p: T1 vs T2	0.156	0.492	0.124
p: T1 vs T3	0.338	0.801	0.270
p: T2 vs T3	0.671	0.360	0.00972
q: T1	0.902	0.902	0.902
q: T2	1	1	1
q: T3	0.578	0.324	0.283

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

Notes: Base category is within-village controls. Robust standard errors in parentheses. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.40: Endline business intention index components: effects by peer type

VARIABLES	(1) Ready to Invest	(2) Submit Plan	(3) Business Likelihood
Trained alone	0.0688	0.0346	0.0825
	(0.0839)	(0.0474)	(0.134)
Close x More Central	0.157	0.0788	0.259
	(0.130)	(0.0755)	(0.191)
Other Pair Types	0.00297	0.0632	0.0552
	(0.0838)	(0.0452)	(0.128)
Degree	0.0366***	0.0203***	0.0247
	(0.0129)	(0.00685)	(0.0207)
Constant	1.144***	0.610***	3.467***
	(0.0879)	(0.0499)	(0.143)
Observations	696	671	555
R-squared	0.022	0.024	0.006
p: Close+Central vs Other	0.235	0.832	0.274
p: Close+Central vs T1	0.496	0.555	0.355
q: Close+Central	0.424	0.424	0.424
q: Other Pairs	1	0.947	1

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

Notes: Base category is within-village controls. "Close x Central" indicates pairing with a socially close peer who is more central. Robust standard errors in parentheses.

Table SA.41: Endline steps index components: ITT effects

VARIABLES	(1) Savings	(2) Loans
Trained alone	0.0690*	0.00349
	(0.0389)	(0.0432)
Treatment with Peer	-0.00258	-0.0758*
	(0.0422)	(0.0456)
Treatment with Peer + Connections Module	0.0360	0.0152
	(0.0420)	(0.0441)
Constant	0.808***	0.797***
	(0.0301)	(0.0308)
Observations	685	680
R-squared	0.006	0.008
p: T1 vs T2	0.0635	0.0803
p: T1 vs T3	0.390	0.790
p: T2 vs T3	0.354	0.0490
q: T1	0.182	0.880
q: T2	0.907	0.240
q: T3	1	1

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

Notes: Base category is within-village controls. Robust standard errors in parentheses. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.42: Endline steps index components: effects by peer type

VARIABLES	(1) Savings	(2) Loans
Trained alone	0.0620	-0.00172
	(0.0388)	(0.0434)
Close x More Central	0.0817	-0.0119
	(0.0604)	(0.0743)
Other Pair Types	0.0376	0.0125
	(0.0392)	(0.0415)
Degree	0.0124**	0.00826
	(0.00561)	(0.00664)
Constant	0.749***	0.757***
	(0.0424)	(0.0454)
Observations	675	671
R-squared	0.024	0.019
p: Close+Central vs Other	0.450	0.739
p: Close+Central vs T1	0.735	0.891
q: Close+Central	0.546	0.775
q: Other Pairs	1	1

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

Notes: Base category is within-village controls. “Close x Central” indicates pairing with a socially close peer who is more central. Robust standard errors in parentheses.

Table SA.43: Endline take-up index components: ITT effects

VARIABLES	(1) Mentoring	(2) Advice	(3) Trainer	(4) Trainer Payment	(5) Mentoring Payment
Trained alone	0.0475	0.0818***	0.0895**	69.78	169.4
	(0.0322)	(0.0314)	(0.0369)	(175.5)	(132.4)
Treatment with Peer	-0.000659	0.0280	0.0405	169.5	90.97
	(0.0352)	(0.0350)	(0.0396)	(170.9)	(105.6)
Treatment with Peer + Connections Module	0.0103	0.0610*	0.0486	138.7	43.02
	(0.0359)	(0.0338)	(0.0406)	(186.8)	(110.4)
Constant	0.874***	0.862***	0.815***	609.8***	388.4***
	(0.0251)	(0.0262)	(0.0296)	(112.4)	(68.84)
Observations	692	690	685	678	688
R-squared	0.004	0.011	0.009	0.001	0.003
p: T1 vs T2	0.131	0.0643	0.154	0.593	0.572
p: T1 vs T3	0.255	0.451	0.250	0.732	0.374
p: T2 vs T3	0.758	0.297	0.833	0.876	0.684
q: T1	0.164	0.0410	0.0410	0.336	0.178
q: T2	1	1	1	1	1
q: T3	1	0.561	0.861	1	1

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

Notes: Base category is within-village controls. Robust standard errors in parentheses. Payment variables are winsorized at the 99th percentile. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.44: Endline take-up index components: effects by peer type

VARIABLES	(1) Mentoring	(2) Advice	(3) Trainer	(4) Trainer Payment	(5) Mentoring Payment
Trained alone	0.0398	0.0773**	0.0792**	12.92	112.5
	(0.0321)	(0.0311)	(0.0365)	(164.3)	(121.2)
Close x More Central	0.101***	0.139***	0.134***	41.53	114.6
	(0.0358)	(0.0263)	(0.0462)	(190.6)	(165.3)
Other Pair Types	0.0302	0.0622**	0.0689*	254.4	87.07
	(0.0319)	(0.0313)	(0.0358)	(176.5)	(101.9)
Degree	0.0155***	0.00871*	0.0198***	-16.87	7.237
	(0.00412)	(0.00496)	(0.00540)	(24.82)	(14.54)
Constant	0.800***	0.820***	0.720***	690.6***	353.7***
	(0.0345)	(0.0382)	(0.0428)	(178.8)	(105.6)
Observations	683	681	677	671	679
R-squared	0.048	0.035	0.052	0.005	0.002
p: Close+Central vs Other	0.0306	3.90e-05	0.120	0.309	0.870
p: Close+Central vs T1	0.0624	0.000583	0.195	0.885	0.991
q: Close+Central	0.00700	0.00100	0.00700	0.495	0.324
q: Other Pairs	0.309	0.159	0.159	0.177	0.309

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

Notes: Base category is within-village controls. "Close x Central" indicates pairing with a socially close peer who is more central. Robust standard errors in parentheses.

Table SA.45: Endline take-up index components: effects by social distance and peer centrality

VARIABLES	(1) Mentoring	(2) Advice	(3) Trainer	(4) Trainer Payment	(5) Mentoring Payment
Trained alone	0.0475	0.0818***	0.0895**	69.78	169.4
	(0.0322)	(0.0315)	(0.0370)	(175.8)	(132.6)
Trained with Peer	0.0236	-0.0618	0.0627	-96.70	80.45
	(0.0561)	(0.0676)	(0.0568)	(428.0)	(285.7)
Trained with Peer × Peer distance	-0.00499	0.0165**	-0.00627	82.35	0.988
	(0.00870)	(0.00784)	(0.0103)	(71.39)	(36.88)
Trained with Peer × Peer degree	0.00895*	0.0169***	0.0101*	7.562	4.138
	(0.00523)	(0.00607)	(0.00553)	(50.76)	(32.75)
Constant	0.874***	0.862***	0.815***	609.8***	388.4***
	(0.0252)	(0.0263)	(0.0297)	(112.6)	(68.95)
Observations	678	676	672	666	674
R-squared	0.037	0.037	0.037	0.005	0.003
q: Alone	0.165	0.0420	0.0420	0.338	0.179
q: Paired	1	1	1	1	1
q: Paired x Dist	1	0.219	1	0.994	1
q: Paired x Deg	0.133	0.0290	0.133	0.562	0.562

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

Notes: Base category is within-village controls. Regression includes interactions of paired treatment with peer social distance and peer degree centrality. Robust standard errors in parentheses. Payment variables are winsorized at the 99th percentile. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

M Follow-up Index Components

SA.1 Outcome Descriptions

The following indices capture outcomes measured one year after the training program via phone surveys.

Outcomes Index. This index captures realized economic outcomes:

1. Monthly income (NPR, winsorized at 1st and 99th percentiles)
2. Agricultural profits (NPR, winsorized at 1st and 99th percentiles)

Steps Index. This index captures concrete actions taken toward starting or expanding a business:

1. Opened a new business in the past year (binary)
2. Investment in new business (NPR, winsorized at 99th percentile)
3. Investment in agriculture (NPR, winsorized at 1st and 99th percentiles)
4. Opened a new savings account with a bank (binary)
5. Joined a new cooperative in the past year (binary)
6. Total savings (NPR, winsorized at 1st and 99th percentiles)
7. Loan amount (NPR, winsorized at 99th percentile)

Mindset Index. This index captures forward-looking motivation and self-beliefs:

1. Income aspirations: "How much would you want to earn per month?" (NPR, winsorized at 1st and 99th percentiles)
2. Plan to start a business (binary)
3. Business self-efficacy (scale 1–5)
4. Committed to a savings account where funds can only be used for business (binary)
5. Amount committed to business savings account (NPR)

Business Practices Index. This index captures improvements in business organization:

1. Keeps physical records of agricultural profits (binary)
2. Keeps physical records of business profits (binary)

Table SA.46: Follow-up outcomes index components: ITT effects

VARIABLES	(1) Monthly Income	(2) Agriculture Profits
Spillover	5,094	31,256
	(4,318)	(30,268)
Trained Alone	4,096	16,310
	(3,626)	(20,086)
Trained with Peer	2,327	3,650
	(3,002)	(19,807)
Constant	23,007***	112,699***
	(2,453)	(14,538)
Observations	538	552
R-squared	0.005	0.006
p: T1 vs Spillover	0.828	0.525
p: T1 vs Paired	0.492	0.287
p: Paired vs Spillover	0.490	0.221
q: Spillover	0.452	0.452
q: T1	0.737	0.737
q: Paired	1	1

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

Notes: Base category is pure control. T2 and T3 are pooled into a single paired treatment indicator. Standard errors clustered at village level. Continuous variables are winsorized at the 1st and 99th percentiles. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.47: Follow-up outcomes index components: effects by peer type

VARIABLES	(1) Monthly Income	(2) Agriculture Profits
Spillover	5,154	32,999
	(4,338)	(30,728)
Trained Alone	3,917	15,675
	(3,623)	(20,932)
Close x Less Central	8,302	38,540
	(6,805)	(28,982)
Other Pair Types	1,232	-6,073
	(3,474)	(22,588)
Degree	65.12	2,287
	(484.2)	(2,686)
Constant	22,636***	100,030***
	(3,188)	(20,898)
Observations	528	543
R-squared	0.009	0.012
p: Close+Less Central vs Other	0.339	0.0466
p: Close+Less Central vs T1	0.522	0.263
q: Close+Less Central	0.305	0.305
q: Other Pairs	1	1

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

Notes: Base category is pure control. "Close x Less Central" indicates pairing with a socially close peer who is less central. Standard errors clustered at village level.

Table SA.48: Follow-up steps index components: ITT effects

VARIABLES	(1) Opened New Business	(2) Investment in New Business	(3) Investment in Agriculture	(4) New Savings Account	(5) Joined Cooperative	(6) Savings	(7) Loan
Spillover	0.0429	3,143	45,430*	0.105*	0.0786	1,937*	23,119
	(0.0295)	(2,844)	(26,014)	(0.0517)	(0.0533)	(973.8)	(17,193)
Trained Alone	0.0188	1,787	27,196	0.0991**	0.112**	1,282	12,060
	(0.0172)	(1,510)	(21,582)	(0.0430)	(0.0478)	(814.2)	(14,915)
Trained with Peer	0.0139	1,091	18,786	0.0304	0.00104	952.9	16,182
	(0.0114)	(936.1)	(15,492)	(0.0358)	(0.0385)	(575.8)	(14,759)
Constant	0.0143*	857.1	98,953***	0.171***	0.0929***	2,177***	21,071**
	(0.00706)	(603.9)	(9,898)	(0.0205)	(0.0329)	(361.5)	(9,752)
Observations	579	579	435	580	580	527	580
R-squared	0.006	0.006	0.011	0.010	0.020	0.013	0.003
p: T1 vs Spillover	0.400	0.583	0.513	0.918	0.440	0.549	0.507
p: T1 vs Paired	0.763	0.674	0.622	0.127	0.00736	0.649	0.797
p: Paired vs Spillover	0.254	0.444	0.262	0.190	0.0955	0.269	0.624
q: Spillover	0.251	0.251	0.251	0.251	0.251	0.251	0.251
q: T1	0.314	0.314	0.314	0.115	0.115	0.270	0.424
q: Paired	0.657	0.657	0.657	0.657	0.657	0.657	0.657

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

Notes: Base category is pure control. T2 and T3 are pooled into a single paired treatment indicator. Standard errors clustered at village level. Continuous variables are winsorized at the 1st and 99th percentiles (business investment and loans at the 99th percentile only). Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.49: Follow-up steps index components: effects by peer type

VARIABLES	(1) Opened New Business	(2) Investment in New Business	(3) Investment in Agriculture	(4) New Savings Account	(5) Joined Cooperative	(6) Savings	(7) Loan
Spillover	0.0421	3,043	46,174*	0.111**	0.0776	1,948*	23,034
	(0.0294)	(2,834)	(26,130)	(0.0511)	(0.0532)	(982.5)	(16,982)
Trained Alone	0.0204	2,016	27,142	0.0983**	0.106**	972.8	12,015
	(0.0174)	(1,521)	(21,487)	(0.0410)	(0.0479)	(696.7)	(14,743)
Close x Less Central	-0.00797	104.8	10,209	0.0103	0.0529	1,416	15,592
	(0.00904)	(650.4)	(36,834)	(0.0685)	(0.0706)	(1,382)	(29,971)
Other Pair Types	0.0183	1,554	12,723	0.0133	-0.00402	300.1	13,235
	(0.0150)	(1,224)	(17,838)	(0.0428)	(0.0404)	(491.2)	(16,631)
Degree	-0.00291	-438.5**	1,227	0.00126	-0.00161	126.4	310.4
	(0.00264)	(199.5)	(3,508)	(0.00741)	(0.00504)	(145.2)	(2,572)
Constant	0.0286*	3,012**	92,124***	0.159***	0.101**	1,560**	19,702
	(0.0150)	(1,260)	(19,336)	(0.0456)	(0.0392)	(644.4)	(15,770)
Observations	569	569	428	570	570	517	570
R-squared	0.009	0.011	0.013	0.015	0.019	0.019	0.003
p: Close+Less Central vs Other	0.0478	0.116	0.947	0.967	0.348	0.352	0.940
p: Close+Less Central vs T1	0.105	0.175	0.681	0.260	0.426	0.702	0.891
q: Close+Less Central	1	1	1	1	1	1	1
q: Other Pairs	1	1	1	1	1	1	1

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

Notes: Base category is pure control. "Close x Less Central" indicates pairing with a socially close peer who is less central. Standard errors clustered at village level.

Table SA.50: Follow-up mindset index components: ITT effects

VARIABLES	(1) Income Aspirations	(2) Plan to Start Business	(3) Self-Efficacy	(4) Commitment Savings	(5) Commitment Amount
Spillover	3,043	0.0907	-0.176	0.171**	981.5
	(6,698)	(0.377)	(0.319)	(0.0708)	(819.4)
Trained Alone	-1,495	0.357	0.0302	0.0612	906.2
	(5,302)	(0.377)	(0.326)	(0.0719)	(686.1)
Trained with Peer	-5,909	0.0724	-0.336	0.0541	734.4
	(4,461)	(0.358)	(0.303)	(0.0679)	(619.3)
Constant	52,329***	2.181***	2.871***	0.439***	900.4*
	(4,071)	(0.313)	(0.261)	(0.0524)	(493.5)
Observations	577	569	580	579	579
R-squared	0.008	0.005	0.007	0.012	0.005
p: T1 vs Spillover	0.499	0.175	0.393	0.0763	0.914
p: T1 vs Paired	0.237	0.159	0.117	0.895	0.658
p: Paired vs Spillover	0.106	0.936	0.446	0.0476	0.705
q: Spillover	1	1	1	0.132	0.937
q: T1	1	1	1	1	1
q: Paired	0.857	0.857	0.857	0.857	0.857

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

Notes: Base category is pure control. T2 and T3 are pooled into a single paired treatment indicator. Standard errors clustered at village level. Income aspirations are winsorized at the 1st and 99th percentiles. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.51: Follow-up mindset index components: effects by peer type

VARIABLES	(1) Income Aspirations	(2) Plan to Start Business	(3) Self-Efficacy	(4) Commitment Savings	(5) Commitment Amount
Spillover	2,938	0.104	-0.154	0.177**	1,023
	(6,690)	(0.372)	(0.319)	(0.0703)	(816.5)
Trained Alone	-1,637	0.369	-0.00540	0.0553	839.7
	(5,189)	(0.371)	(0.329)	(0.0695)	(669.9)
Close x Less Central	-2,152	0.495	-0.148	0.0409	1,472
	(9,706)	(0.436)	(0.408)	(0.106)	(1,883)
Other Pair Types	-6,406	0.0904	-0.321	0.0527	378.0
	(4,746)	(0.352)	(0.317)	(0.0690)	(584.1)
Degree	-410.9	0.00144	0.0643	0.0111	146.6
	(606.7)	(0.0394)	(0.0388)	(0.0103)	(109.4)
Constant	54,358***	2.161***	2.548***	0.380***	172.0
	(4,388)	(0.377)	(0.332)	(0.0831)	(697.5)
Observations	567	560	570	569	569
R-squared	0.009	0.008	0.014	0.016	0.012
p: Close+Less Central vs Other	0.652	0.241	0.590	0.905	0.555
p: Close+Less Central vs T1	0.956	0.702	0.707	0.890	0.739
q: Close+Less Central	1	1	1	1	1
q: Other Pairs	1	1	1	1	1

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

Notes: Base category is pure control. "Close x Less Central" indicates pairing with a socially close peer who is less central. Standard errors clustered at village level.

Table SA.52: Follow-up business practices components: ITT effects

VARIABLES	(1) Keeps Agriculture Records	(2) Keeps Business Records
Spillover	-0.00476	0.107
	(0.0526)	(0.0749)
Trained Alone	0.00913	-0.0334
	(0.0594)	(0.0730)
Trained with Peer	-0.0572	0.0400
	(0.0427)	(0.0636)
Constant	0.729***	0.331***
	(0.0250)	(0.0511)
Observations	580	578
R-squared	0.004	0.009
p: T1 vs Spillover	0.844	0.0500
p: T1 vs Paired	0.213	0.155
p: Paired vs Spillover	0.349	0.242
q: Spillover	0.867	0.492
q: T1	1	1
q: Paired	0.623	0.623

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

Notes: Base category is pure control. T2 and T3 are pooled into a single paired treatment indicator. Standard errors clustered at village level. Q-values are FDR-adjusted using the Benjamini et al. (2006) sharpened procedure, computed separately for each treatment coefficient across outcomes within each index family.

Table SA.53: Follow-up business practices components: effects by peer type

VARIABLES	(1) Keeps Agriculture Records	(2) Keeps Business Records
Spillover	-0.000995	0.111
	(0.0529)	(0.0762)
Trained Alone	0.00434	-0.0315
	(0.0604)	(0.0736)
Close x Less Central	-0.0311	0.114
	(0.0717)	(0.0974)
Other Pair Types	-0.0278	0.0318
	(0.0432)	(0.0715)
Degree	0.00849	-0.00538
	(0.0106)	(0.0117)
Constant	0.685***	0.353***
	(0.0601)	(0.0762)
Observations	570	568
R-squared	0.006	0.013
p: Close+Less Central vs Other	0.960	0.422
p: Close+Less Central vs T1	0.544	0.123
q: Close+Less Central	1	1
q: Other Pairs	1	1

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

Notes: Base category is pure control. "Close x Less Central" indicates pairing with a socially close peer who is less central. Standard errors clustered at village level.