The Value of Strategic Networking [†]

Vatsal Khandelwal [§]

Juni Singh [¶]

January 22, 2025

Click here to see the most recent version.

Abstract

How important is it to network strategically? We implement a RCT in rural Nepal to study the effects of entrepreneurship training and contact-sharing among women who differ in terms of their closeness and connectedness in their social network. The intervention improves short-term outcomes when individuals are paired with a more connected, socially close peer and long-term outcomes when paired with a less connected, socially close peer. These peer effects arise due to motivation in the short term and ease of collaboration in the long term. Counterfactual exercises show that interacting strategically can improve outcomes by 0.8-1.1σ compared to random interactions.

KEYWORDS: Social Networks, Peer Effects, Networking, Entrepreneurship, RCT.
JEL CLASSIFICATION: D85, O12, Z13, J16, L14, L26.

[†]We thank the Weiss Fund for providing funding to conduct the fieldwork. We also thank our partner organizations, JPAL and Governance Lab, and our enumerators for the execution of the field project. We are very grateful to Marina Agranov, Abhijit Banerjee, Emily Breza, Michael Callen, Margherita Comola, Horacio Larreguy, Sylvie Lambert, Karen Macours, 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 has 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

Strategic networking is frequently cited as a determinant of success- from choosing whom to sit next to at a conference to identifying a helpful mentor.[1] While the economic benefits of social networks are well documented across various domains[2], it is unclear if it is beneficial to have strategic interactions with peers, especially when it may require costly information about the structure of the network. For example, while interacting with popular peers may boost motivation and interacting with socially close peers may encourage collaboration, it is an open question whether the benefits of interacting with a strategy outweigh the benefits of interacting at random. It is also unclear whether networking strategies should adapt to different goals or evolve over time. Answering these questions is challenging due to the lack of exogenous variation in the structure of social networks.

In this paper, we induce exogenous interactions in social networks to show that strategic networking can improve short term and long term outcomes around female entrepreneurship. While existing research shows that networks can help improve business outcomes (Field et al. 2016, Carranza et al. 2018, Asiedu et al. 2023), we ask if specific types of peers are better suited to providing specific types of support. We leverage random variation in peer identity and data from multiple time periods to identify heterogeneous peer effects and show that strategic networking is beneficial both in the short and long-term. We find that the benefits of networking are primarily direct, i.e. individuals derive a direct benefit from interacting with peers as opposed to an indirect benefit by using peers to access the wider social network. We then show that networking strategies must depend on whether we intend to improve short-term or long-term outcomes. Finally, counterfactual exercises demonstrate that policymakers can feasibly engineer such strategic interactions to generate large benefits.

To do this, we measure the social networks of 2800 women and then implement a randomized controlled trial with 1200 women across 29 villages in rural Nepal. 85% of our sample aspire to a higher income, 42% express a willingness to open businesses, but 50% do not think they have the skills to do so. We implemented an intensive three-day entrepreneurship training program which combined the ILO's standard training curriculum with elements of personal initiative training that focus on raising aspirations.[3] Our experiment design allows us to study how peer interactions can strengthen the effects of this core program. While individuals in pure control villages did not attend the training, individuals in treated

[1] For example, see the Forbes article "How Leaders Strategically Enhance Their Networks For Maximum Impact" (Cecchi-DiMeglio 2023) on the benefits of strategic networking in a professional context.
[2] See Breza (2016), Breza et al. (2019) for a review of this literature.
[3] McKenzie et al. (2021) review the effects of both types of training programs as implemented in existing literature, the majority of which, unlike us, focus on individuals who already own businesses.

villages were allocated into one of three groups. They either participated in the training and completed all training activities alone, attended the training with a *randomly* chosen woman in the village, or did not attend it at all. Among those who attended in randomly formed pairs, we additionally randomized the implementation of a "connections module" in which randomly paired trainees were asked to enlist their network contacts together and think of ways these contacts can help them in setting up a business.

We leverage the design to answer the following questions. First, we exploit the variation in whether women attend the training alone versus with a randomly chosen person in their social network to assess whether social support can improve outcomes. We expect that peers can provide support through skill complementarities, motivation, and risk-sharing.[4] Moreover, social networks in this setting are sparse and exhibit a strong community structure suggesting that networking opportunities may otherwise be limited.[5] [6] Second, we leverage whether individuals attended the connections module to compare the direct value of networking (in terms of being trained with peers) with the indirect value (in terms of being able to access their social contacts). Finally, we leverage the random variation in the identity of the peer to study if the treatment effect differs as a function of their network position. In particular, we exploit the random variation in social distance (i.e., how socially close the peer is) and centrality of the peer (i.e., how connected the peer is) to identify the benefits of networking strategically. Heterogeneity along these dimensions was pre-registered in our analysis plan.[7] We hypothesize that while social support decreases with social distance, the effect of centrality is more nuanced. For example, being connected to a central person may provide access to a wider social network but they may be difficult to collaborate with.

We test the effect of the various interventions by measuring short-term outcomes immediately after the training and long-term outcomes after one year. First, we find that in the short term, training significantly improves outcomes across the board. These outcomes include readiness to invest in a business, steps to open a business such as willingness to take a loan or open a savings account, and take-up of additional mentoring and assistance. Second, we find that while the connections module has a higher treatment effect in terms of magnitude, the additional impact is not significant in most cases. This is true even though women

[4] These factors affect the opening of businesses in our baseline data. Those who have opened up businesses in this setting are younger, have received more education, have higher income aspirations, and have lower risk aversion. In fact, a Lasso regression reveals age, risk-aversion, and education as the most relevant correlates of willingness to open a business compared to various other demographic characteristics.

[5] 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).

[6] Links with individuals of the same caste are also 19 percentage points higher, on average than the proportion of same-caste members in the village population indicating high levels of homophily.

[7] Details on how we depart from our pre-analysis plan are provided in Section 3.

in this treatment arm share about six network contacts with each other on average, with significantly higher contact sharing among those with a larger baseline difference in their number of network connections. At the same time, individuals are significantly less likely to provide access to contacts to those in a different caste, implying that the indirect benefits of networking cannot be easily leveraged in settings with entrenched social structures.

Third, we find that being paired does not significantly improve immediate outcomes on average, when compared to being trained alone, except for certain pair types depending on their network position. Interestingly, pairing is always beneficial when the matched person is a friend.[8] When we interact this with the 'centrality' of the peer i.e. how connected they are in the network, we find that pairing has a significant additional effect relative to being trained alone only when individuals are paired with a friend who is more central than them. Individuals paired with a more central friend report an 11% higher willingness to invest in a new business (significant at 10%) relative to those trained alone. Moreover, on average, their take-up index is 0.13 standard deviations higher than those trained alone (significant at 5%) implying that they are more likely to demand additional resources to help with setting up a business.[9] Finally, we find that those trained in a pair are 12% more likely to report wanting to open an agro-business relative to other businesses when compared to those trained alone.[10]

Next, we implemented a follow-up survey with a sub-sample one year after the intervention. We find that only 3% have opened up a new business but the training increases the probability of taking steps to open a new business. Those who were trained are significantly more likely to have saved more, opened up a new savings account, and taken a loan compared to those in pure control villages. 38% of those in the paired treatment arms report having spoken to their peer over the last year to ask for advice, borrow/lend money, or about setting up a business. Further, we find that those who were trained with a peer have a 19% higher monthly income (significant at 10%) level when compared to the pure control group. Importantly, this effect is driven by them having invested more and earned more from agriculture-related activities as opposed to non-agricultural businesses. In particular, those trained with a peer have significantly higher agricultural investments than the pure control group and significantly higher agricultural profits when compared to those trained alone. Importantly, while being paired with a friend is beneficial both in the short and long term, we find that, unlike the short-term results, the positive effects on outcomes in the long term are concentrated among

[8] We define a "friend" as a person who has a network distance less than or equal to 2 i.e. strictly lower than the median social distance equal to 3.

[9] We show that these network effects persist even after controlling for the similarity between the peers along other demographic characteristics such as age, caste, education, marital status, and income.

[10] The reciprocity-based labor provision agreements unique to women in Nepalese villages (Messerschmidt 1981, Sherpa 2005, Bhattarai 2006) may help explain why pairs are more willing to open agro-businesses together.

those paired with a friend who was less central than them.

The benefit from networking with specific types of peers can arise due to multiple mechanisms such as better learning, encouragement, risk sharing, and future collaboration. We show that the effects are primarily driven by motivation from a more central, socially close peer in the short term and the ability to collaborate with a less central, socially close peer in the long term. Various findings, including participant reports, effects on self-efficacy, and the role of individuals with a high in-degree (i.e., those who are listed as friends by several others), suggest that endline outcomes improve due to increased motivation.[11] The reversal of the effects of centrality, in the long term, could be linked to their ability to collaborate. While 38% of those trained in a pair reported reaching out to each other for advice, borrowing-lending money, and discussing ideas about forming a business, those matched with a more central person are 9% less likely to talk about financial matters as opposed to taking advice, although the effect is not significant. At the same time, randomly matched with peers of different castes are significantly less likely to report having spoken over the year. Socially close pairs are also significantly more likely to state a willingness to start a business together than non-friends during the endline. Combined, this evidence suggests that individuals might be less likely to interact with others of a different social status than them after the training.

We reconcile these results using a simple effort-choice model in which an agent's choice of entrepreneurial effort depends on their own characteristics, peer characteristics, and the effort exerted by the peer. Importantly, peer effort can have heterogeneous effects depending on social distance and peer centrality. Taking the model to the data for the observed pairs in Treatments 2 and 3, we first find correlational evidence showing that peer centrality does indeed increase the magnitude of peer effect in the short term, and social closeness increases it in the long term. This suggests that strategic complementarities in effort can be higher for pairs with different network configurations and that such heterogeneity may vary over time. Then, we estimate peer effects using data from two survey waves to address the reflection problem (Manski 1993), which can otherwise bias our estimates despite random matching. We find that the effect of the peer's endline entrepreneurial effort on follow-up outcomes is significantly higher (at 5%) for those who are matched with peers at a lower social distance.[12]

Finally, we present the results of policy counterfactuals by leveraging our experimental design to estimate dyadic regressions. We use these regressions to compare the effects of network-

[11] For instance, the majority of paired participants in the endline revealed that motivation was the reason why they thought pairing was beneficial – with this being more likely to be stated as a reason by those who had lower income aspirations in the baseline.

[12] The effect persists even after controlling for the individual's and peer's characteristics and the individual's own outcome in the endline.

based pairing with random pairing. We do this by comparing random pairing with two alternative scenarios: pairing with a socially close individual and pairing with an individual whose network centrality is one standard deviation larger. We find that relative to random pairing, forming pairs by leveraging the characteristics of the underlying social networks (in terms of social distance and differences in centralities) can improve the average outcomes of the pair by 0.8 and 1.1 standard deviations in the short and long-run respectively. While pairings based on social distance can be easy to implement, the feasibility of pairing based on centralities, given the fixed social network, remains unclear. We simulate 10,000 counterfactual reassignments of peer pairings within each village in our dataset and find that strategic pairings based on centralities are indeed feasible and can generate high returns. Having said that, it is important to note that we cannot comment on the private costs of networking that might naturally arise in the absence of policy support. These range from informational costs (eg: knowing who is central) to the costs of approaching others (eg: confidence, time). While we cannot comment on the latter, we find that the majority of individuals in our setting guess the number of connections of their randomly matched peer within a range of $\pm$ 2 and about 41% guess it correctly within a range of $\pm$ 1.[13] This suggests that information costs might be low. Setting up platforms to minimize interaction costs can then be a way to leverage the benefits of strategic networking.

Our paper contributes to the existing literature in three ways.

First, we contribute to the large literature on peer effects in economics. Peer effects have been measured in various contexts, from 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 that estimates peer effects employs a linear-in-means model (Manski 1993) largely treating peers as a homogeneous group whose mean choices affect the choices of the individual.[14] Recent literature leverages random 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). Contributing to this growing literature, we show that the effects of networking with individuals in distinct network positions do differ in magnitude and also change over time. If the main aim is to provide short-term motivation, then networking with central peers might

[13]This aligns with findings in the literature that show that individuals can possess accurate knowledge about who is central in their network (Banerjee et al. 2019).

[14]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 this assumption can be misleading.

be helpful. However, if the main aim is to form long-term collaborations then networking with less central peers might be sensible. We also find that leveraging this heterogeneity to develop an adaptable networking strategy is both feasible and can generate high returns.

Second, we contribute to a large 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, Carranza et al. 2018, Campos et al. 2019, Asiedu et al. 2023, Vega-Redondo et al. 2023). McKenzie et al. (2021) provide a comprehensive review of the literature on training programs. Unlike most of the existing literature, we focus on a sample of non-business owners and examine both the immediate and one-year impacts of business training on this group. We find that a collaborative business training program with a peer component can lead individuals to expand existing agricultural activities. 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. By doing so, we contribute to existing literature that documents the average treatment effect of peers in improving business or agricultural outcomes (Field et al. 2016, Cai & Szeidl 2018, Vasilaky & Leonard 2018). Moreover, we show that peer effects can operate via different mechanisms over time and that peer matches must be designed accordingly to achieve the desired objective. Related to our paper, Vega-Redondo et al. (2023) also 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 examine the one-to-one impact of a peer's network identity within the same interaction format, focusing on how individuals in the same social network and cultural context can support each other and how this support varies over time.

Finally, our experimental design allows us to distinguish between two contrasting mechanisms through which networking can be beneficial. Peers can directly influence outcomes, or they can serve as an indirect gateway to a broader social network. We disentangle these effects using a novel 'connections module'. This design enables us to assess whether the indirect effect is empirically significant, contributing to the literature on "bonding" and "bridging" social capital first introduced by (Putnam 2000). Our findings that the indirect channel does not produce meaningful effects, despite significant sharing of contacts, suggest that a lack of information about how network members can support business development may not be a binding constraint in this context.

The paper is organized as follows. Section 2 describes the context and experiment design. We discuss the 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 provide a theoretical framework, identify peer effects, and present policy counterfactuals. Section 7 concludes.

2 Data and Experiment Design

We conducted our surveys and experiments in 29 villages in rural Nepal. This includes a baseline survey with about 2800 women in 2021, the randomised controlled trial and endline survey with 1200 women in 2022, and a follow-up survey with about 750 women in 2023.

We first conducted a detailed baseline survey with about 2800 women across all villages in our sample. We collected data on various demographic characteristics along with data on social networks. Additionally, we also measured aspirations around agricultural investments, non-agricultural investments, and income.[15] On average, a village contains 70 households, with an average of 100 women per village. We used a village census to administer the network survey. The networks questionnaire included questions designed to elicit information about social networks, as in Banerjee et al. (2013). These questions measure whom individuals report interacting with in any capacity including visiting homes, giving or taking advice, going to a temple with, or contacting during a health emergency. Individuals were asked to list as many names as they preferred. The links are undirected in that i is assumed to be a friend of j if either of them mentions each other's names for any of the above interactions.

2.1 Baseline Findings

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 46% of our sample have no formal education. We find that roughly 22% of women report having opened businesses already but 42% report a willingness to open a non-agricultural business. 85% aspire to earn an income higher than their current income while 25% aspire to spend more on non-agricultural business expenditures than their current investment. 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.[16]

[15] Data on aspirations were collected in line with the procedure outlined in Bernard & Seyoum Taffesse (2014).
[16] Risk preferences were elicited using a choice experiment involving a series of lotteries and a fixed payment.

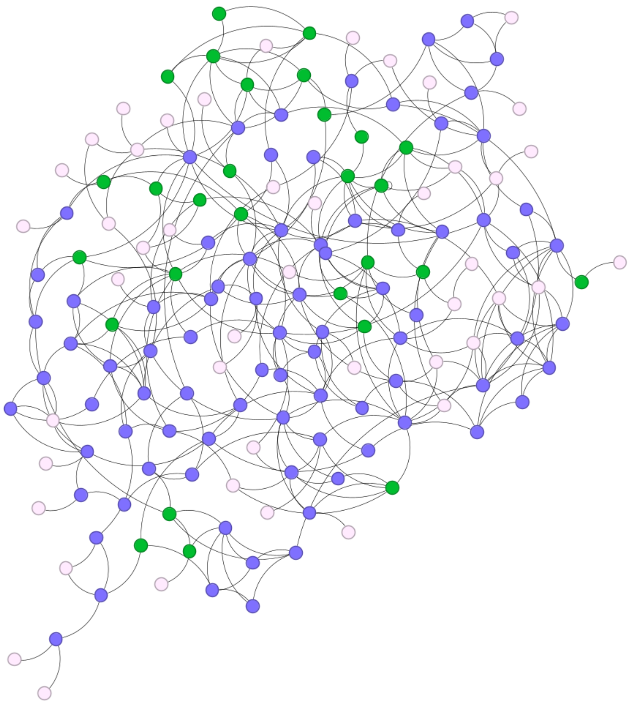


Figure 1: Graph of the social network in an example village. Surveyed women with and without an existing business are coloured in green and blue respectively.

2.1.1 Network Statistics

The average number of connections in the undirected friendship network is ~ 5 links. Table A.2 shows that the number of connections is correlated with other demographic variables of interest including education, caste, and marital status. Figure 2 plots the density of networks in various villages. 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 in each network. This suggests that the networks are very sparse.

As shown in Figure 3, these networks also exhibit a community structure and a high level of homophily: intra-caste links are more common than links across castes. On average, the proportion of same-caste links exceeds the proportion of same-caste village members by 19 percentage points. These findings suggest that pairing individuals at random and having them interact would involve networking that does not already occur in this setting.

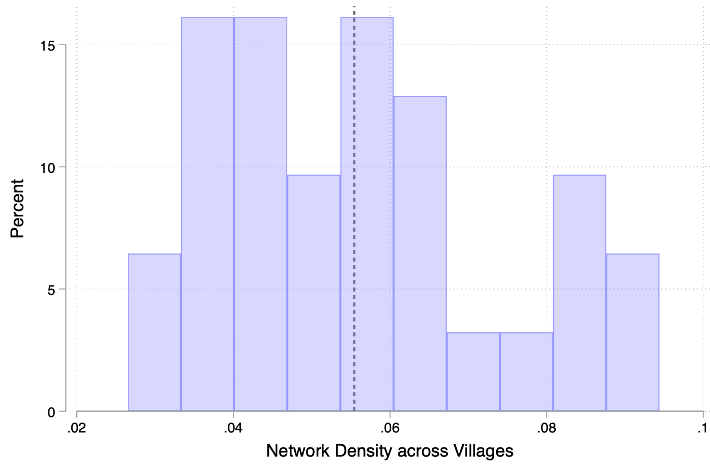


Figure 2: The above histogram plots the network density for each village in the sample. Density is computed as the number of realized edges as a proportion of total possible edges. The number of total possible edges is given by $\frac{n(n-1)}{2}$ where n is the number of individuals in the network.

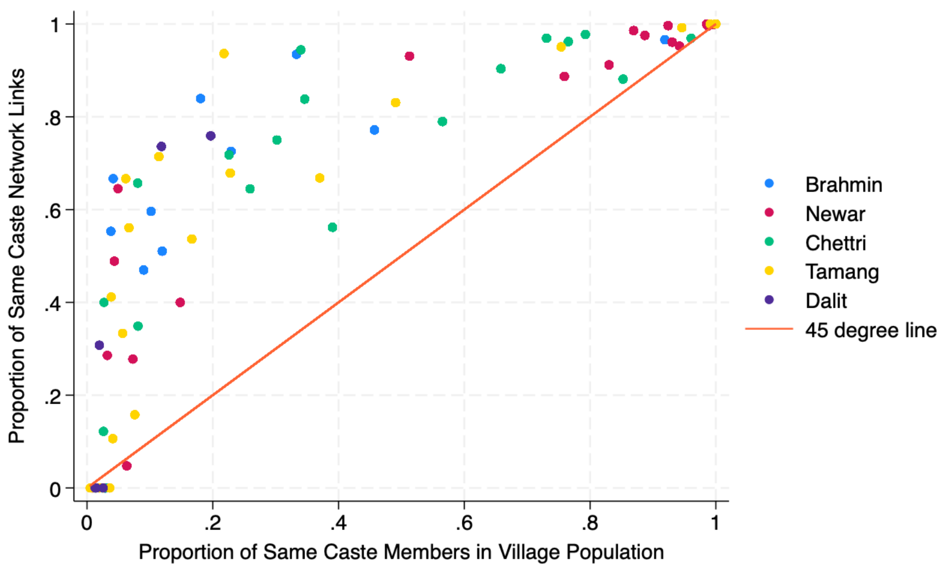


Figure 3: This graph plots the proportion of network links between individuals of the same caste against the proportion of same caste individuals in the village. Each dot represents a village-caste combination. As the graph shows, the proportion of same-caste network links consistently exceeds the proportion of same-caste individuals in the village, indicating a pattern of homophily in linking behavior.

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

We find that women who already own a business have a 20% higher income than the ones with no business ownership. We correlate whether or not they have opened businesses with their baseline characteristics. These results are shown in Table A.3. 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 4 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 those who haven't opened up businesses were asked about why they haven't done so already, we find that 50% say that do not have the skills, 28% feel they are not capable, and 23% say that they lack the financial ability. We correlate 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 A.4. Figure 5 shows the results of a Lasso regression which selects age, risk-aversion, and education as the relevant correlates at high values of the penalty parameter. We find that those who are older and more risk-averse are correlated with being less willing to open businesses. This suggests that risk-sharing with peers could assist in opening up businesses. Moreover, we find that those who are more educated are correlated with being more likely to open businesses suggesting that skill complementarities with peers might be helpful as well. Finally, those who have higher aspirations are more willing to open businesses suggesting that peers can potentially be used to motivate and boost aspirations that can then be channelled into opening businesses.

2.2 Experiment Design

We conducted our experiment in September 2022.

The experiment consisted of a three-day entrepreneurship training motivated by the SIYB module developed by the International Labour Organisation. The training typically lasted 3 hours per day and individuals were given 100 Rs/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. We largely excluded women who already own a non-agricultural business ensuring that the majority of our experiment sample comprises of women who have never opened a business before.[17]

Figure 6 represents our two-step randomisation design. First, we allocated villages to Pure

[17]About 97% of our endline sample includes women whose main source of income is agriculture. Out of these, about 20% report their main source of income to be an agriculture-related business.

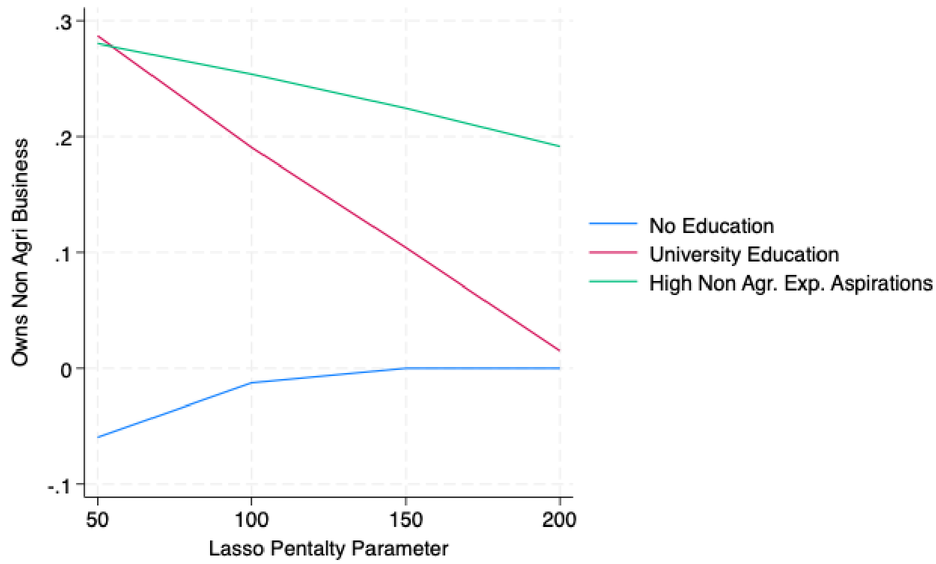


Figure 4: The above graph plots correlations between selected variables and owning a business for different values of penalty parameters in a Lasso regression. Only variables selected out of a large list of demographic and network variables at $\lambda = 100$ are displayed. No Education and University 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.

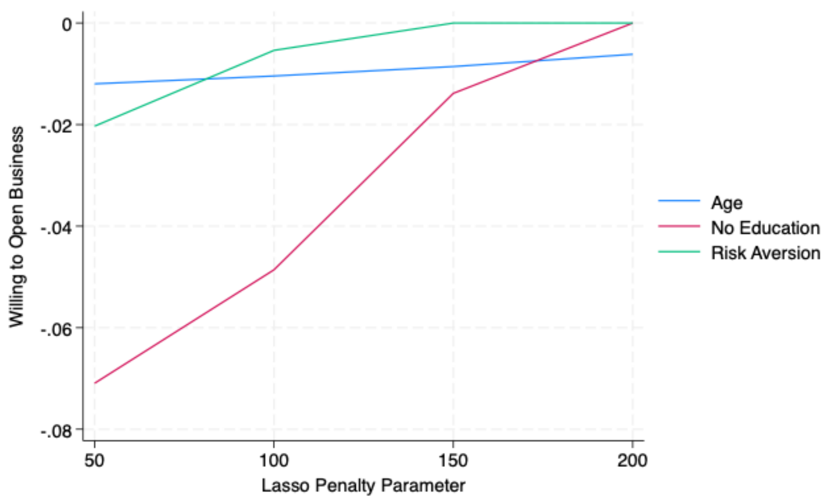


Figure 5: The above graph plots correlations between selected variables and willingness to open businesses for different values of penalty parameters in a Lasso regression. Only variables selected out of a large list of demographic and network variables at $\lambda = 100$ are displayed. No Education 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.

Control and Treatment. Those in treatment villages were then randomly allocated either to the control group or one of the three treatment arms across all villages at an individual level. Women in treated villages were randomly allocated into one of the four groups:

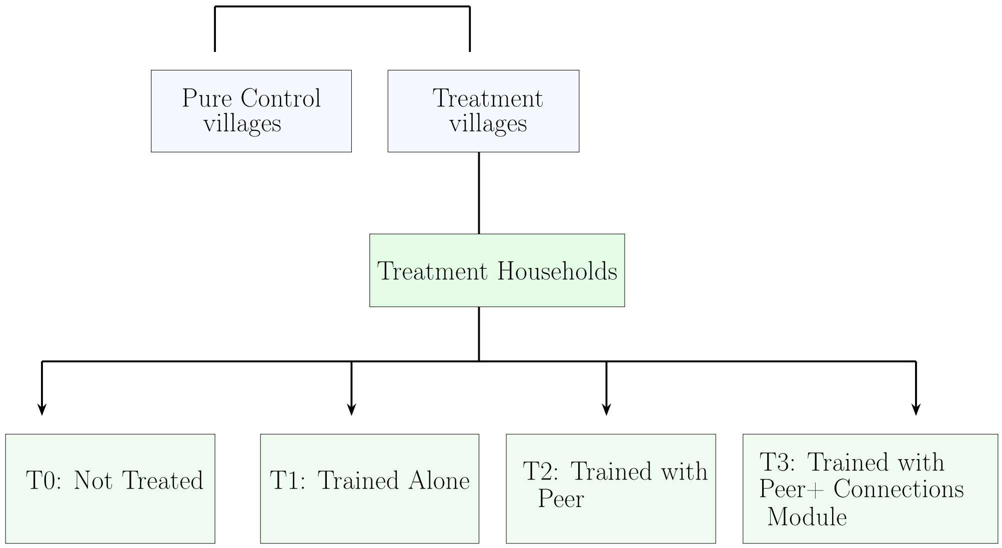


Figure 6: Experiment Design

- *Control*: Women in this group did not receive training but live in Treatment villages. They are treated as the spillover group in the one-year follow-up.
- *Treatment 1*: Women in this group attended the training alone.
- *Treatment 2*: Women in this group attended the training with another randomly chosen person from their social network. The pair varies in how connected each respondent is and the social distance between them.
- *Treatment 3*: Women in this group also attended the training with another randomly chosen person from their social network. In addition, the pair are provided with a 30-minute connections module that encourages 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 can help each other. Following this, the pair were asked to list 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. As before, the pair vary in how central both of them are in the social network and the social distance between them.

3 Estimation Strategy

We measure the impact of the training using various specifications that help us understand if the training is helpful in general, if it has higher returns for those who were paired, if it has higher returns for those who attended the connections module, and if it works for pairs with specific network characteristics.

3.1 Endline Outcomes

We first measure the effect of the treatment on four main endline outcomes immediately after the training. This includes the Business Aspirations Index, Business Index, Take-up Index, and Readiness to Invest. We measure these outcomes using our endline survey. In each case, we construct a weighted average using the weighting procedure proposed in Anderson (2008). Each index is normalised using an indicator variable for the control group. The description of the variables that enter each index is as follows:

(1) **Business Aspirations Index:** We compute a measure of business-related aspirations that comprises of yearly non-agricultural investment, monthly income, and savings. We elicit aspirations by following the procedure in Bernard & Seyoum Taffesse (2014). We ask individuals the minimum and maximum of the relevant variables in their neighbourhood, the current value, and what they aspire to in the corresponding time frame. We construct the weighted average of the individual's aspirations on all of these dimensions to create the index. (2) **Business Index:** For the business index, we construct a weighted average of various variables including whether the individual thinks they have the skills to open a business, on a scale of 1 to 5 how ready they are to start a business, if they are willing to submit their business plan for a hypothetical competition, whether they want to take a loan for the business, and whether they want to open a savings account for their business. These questions elicit willingness to go ahead with these activities. (3) **Takeup Index:** The takeup index is constructed by taking a weighted average of variables that indicate that the individual seeks additional assistance to set up a business. We ask the following questions related to the hypothetical takeup of options. We ask if individuals would take up the opportunity of additional paid trainers and mentoring workshops in the next year. If so, we ask how much they would be willing to pay for each of those opportunities. We also asked if they were willing to take advice from members of their community regarding opening a business. (4) **Readiness to Invest:** Finally, we measure the impact of the treatment on a binary variable indicating whether the respondent reported that they were willing to invest in a business and if so, the kind of business they would like to start.

We also collect information on a variety of other secondary variables ranging from the types of businesses individuals want to open or their intention to follow up with their matched peers. These variables will be defined as we proceed.

3.2 Follow-up Outcomes

We measure outcomes for a random subsample of ~ 750 individuals one year after the training. 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 for their business from where they cannot take out funds unless they use it for business purposes, record-keeping for agriculture, and other outcomes regarding community interactions around advice-taking and collaborations.

3.3 Empirical Specifications

3.3.1 Impact of Training

We measure the impact of the treatments using the main specification described below:

$$Y_i = \alpha + \beta_1 T1_i + \beta_2 T2_i + \beta_3 T3_i + \epsilon_i$$

Y_i is an outcome measure for individual i, $T1_i$ is a dummy variable that takes value 1 if the individual was treated alone and 0 otherwise. Similarly, $T2_i$ is a dummy variable that takes value 1 if the individual was treated with a pair and $T3_i$ is a dummy variable that takes value 1 if the individual was treated with a pair and an additional connection module that emphasizes the importance of networks. Standard errors are clustered at the village level. We additionally control for an indicator variable for the spillover group in any regression run on the follow-up survey sample that measures outcomes one year after the intervention.

3.3.2 Impact of Training with a Peer

In this specification, we club treatments 2 and 3 together and create an indicator variable that is equal to 1 for all women who are treated in a pair. We regress outcome variables on the treatment dummies using the specification described below:

$$Y_i = \alpha + \beta_1 T1_i + \beta_2 (T2_i + T3_i) + \epsilon_i$$

Y_i is an outcome measure for individual i, $T1_i$ is a dummy variable that takes value 1 if the individual was treated alone, $T2_i$+ $T3_i$ is a dummy variable that takes value 1 if the individual was treated with a pair. Standard errors are clustered at the level of the village. As before, for the follow-up sample (measuring outcomes one year after the intervention), we additionally control for an indicator variable for the spillover group.

3.3.3 Impact of Training with different Types of Peers

Next, 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. [18] Let d_{ij} be the network distance between i and j and let ϕ_i be the network centrality (eg: number of connections) of agent i. First, we combine $T2$ and $T3$ and split individuals into whether or not they are socially close. We define an individual as a socially close (henceforth "friend") if the network distance between them is strictly less than the median social distance $d_{ij} = 3$.

Then, we estimate the following regression:

$$Y_{iv} = \alpha + \beta_1 Friend + \beta_2 NonFriend + \epsilon_i$$

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 the respondent and implement a similar specification as above. Following this, we classify peer type into one of four interaction categories: Friend X Central, Friend X Noncentral, Non Friend X Central and Non Friend X Non Central. For any pair ij, i is assigned to the category friend-central if their matched "friend" j has strictly higher degree centrality

[18] Due to concerns around statistical power, we differ from our pre-analysis plan in that we do not use the social distance and centrality of the matched peer as controls in these regressions but instead use binary variables indicating whether they are socially close or whether they are more or less central than the respondent.

compared to i, i.e ϕ_i- $\phi_j < 0$ 0 and i and j have social distance $d_{\ ij<=2}$.

$$Y_{iv} = \alpha + \beta_1 FriendXCentral + \beta_2 FriendXNonCentral \\ + \beta_3 NonFriendXCentral + \beta_4 NonFriendXNonCentral + \epsilon_v$$

We exclude the control group in these specifications so that we can compare the effect of being trained with a specific type of peer to being trained alone.

It is important to note that an individual who is central in the network is mechanically more likely to be paired with a friend or to have a less central peer. This implies that an individual's centrality can be correlated with the group to which they are allocated in the above specifications, even though pairing is implemented at random. To ensure that this does not affect our results, we control for the individual's degree centrality in all regressions that compare being trained with a specific type of peer to being trained alone.

3.3.4 Dyadic Regressions and Linear-in-Means Specifications

In addition to the above, we will also present dyad-level specifications and identify peer effects by exploiting random pairing and data from multiple survey waves to estimate a linear-in-means specification (Manski 1993). These results will be described and presented after the reduced form results in Section 6. These results further strengthen our reduced form results, especially in cases where we lack statistical power.

4 Results

4.1 Balance

Before proceeding with the results, we first check for balance in baseline characteristics among the control and treatment groups. We check for balance on baseline variables 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 across all treatment arms in the endline and follow-up sample.

Table B.1 and B.2 show these results. We find that the sample is balanced on most characteristics across all pair-wise comparisons between treatment arms. In particular, out of 114 comparisons, only 7 are significant for the endline sample, and 8 in the follow-up sample (at the 5% level). Next, we also compare individuals in the pure control villages with individuals in treatment villages, as shown in Table B.3. This is critical, especially for the results of the 1-year follow-up survey where we will separately compare the pure control group with the various treatment groups in treated villages. We find that individuals in the pure control group and various treatment groups are similar in most characteristics. Finally, to ensure that our results are robust to any observed cases of imbalance (eg: degree centrality), we will present a robust version of all the main results by employing post double selection Lasso using the method in Belloni et al. (2014).

4.2 Short Term Effects

In this section, we look at the impact of the training immediately after the end of the third day of training. As seen from Table 1, the treatments do not affect business aspirations. However, Treatment 2 and Treatment 3 significantly increase the individual's stated readiness to invest in a business by 9 and 14 percentage points respectively. The effect is insignificant for Treatment 1 and significantly lower than the effect of Treatment 3 (at 10%).

All treatments led to a significant increase in the business index and take-up index. The differences between Treatment 1, 2, and 3 are not statistically significant for all these outcomes implying (i) that pairing, on average, does not lead to a higher treatment effect, and (ii) the connections module while leading to a treatment effect of higher magnitude for readiness to invest and business index, does not have a significant additional impact. This is despite the fact that individuals assigned to the connections module shared 6 contacts on average with each other. In fact, as shown in Table D.5 in the appendix, more connections were pooled among individuals in the same caste and among those with a higher difference in baseline degrees.

4.2.1 Being trained with a peer is marginally better.

As seen above, we find that on average, being paired does not necessarily help improve the outcomes of the training. Table 2 shows that this is true even when we combine Treatment 2 and 3 into one indicator variable that is equal to 1 if the individual is in either arm. We find that pairing does not lead to significant additional improvement in endline outcomes when compared to Treatment 1, even though the effect is always larger in magnitude than that of not being paired. For example, being paired leads to a 12 percentage point increase (significant at 1%) in readiness to invest compared to a 7 percentage point increase (insignificant) for being treated alone. However, this difference is not statistically significant.

Table 1: Impact of the training on immediate outcomes.

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	-0.0975	0.0680	0.343***	0.288***
	(0.0970)	(0.0474)	(0.0891)	(0.0876)
Treatment 2	-0.0203	0.0940*	0.304***	0.355***
	(0.0822)	(0.0467)	(0.100)	(0.0907)
Treatment 3	0.0154	0.141***	0.392***	0.294***
	(0.0803)	(0.0392)	(0.0780)	(0.0792)
Constant	0	0.704***	0	0
	(0.0578)	(0.0283)	(0.0716)	(0.0659)
Observations	1,204	1,199	1,186	1,183
R-squared	0.002	0.016	0.036	0.035
Treatment 1==2	0.448	0.635	0.699	0.407
Treatment 2==3	0.682	0.432	0.375	0.471
Treatment 1==3	0.183	0.0910	0.454	0.924

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

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

Table 2: Impact of the training on immediate outcomes (Paired v/s Unpaired)

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	-0.0975	0.0680	0.343***	0.288***
	(0.0970)	(0.0474)	(0.0891)	(0.0876)
Treatment 2 and 3	-0.00279	0.117***	0.347***	0.325***
	(0.0688)	(0.0314)	(0.0754)	(0.0743)
Constant	0	0.704***	0	0
	(0.0578)	(0.0283)	(0.0716)	(0.0659)
Observations	1,204	1,199	1,186	1,183
R-squared	0.002	0.015	0.035	0.035
T1= Paired	0.254	0.212	0.955	0.522

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

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

4.2.2 Peers are more likely to want to open agricultural businesses.

Table 3 shows the types of businesses individuals report that they want to open after being trained, conditional on being ready to invest. We find that those who are paired together during the training are significantly more likely to want to open an agricultural business together when compared to the control group.

This result could be in line with the existence of the *parma* or *bola* system in Nepalese villages where peers support each other with agriculture-related tasks through reciprocity-based labour provision programs (Messerschmidt 1981, Sherpa 2005, Bhattarai 2006). We will revisit this finding when we present the results of the survey conducted one year later.

Table 3: Types of Businesses that individuals are willing to open

VARIABLES	(1) Agricultural Business	(2) Sewing	(3) Shop/Parlor	(4) Other Business
Treatment 1	0.0368	0.0370	-0.0923*	0.0184
	(0.0470)	(0.0341)	(0.0494)	(0.0173)
Treatment 2 and 3	0.116**	-0.0269	-0.0956*	0.00681
	(0.0436)	(0.0243)	(0.0471)	(0.0119)
Constant	0.513***	0.116***	0.346***	0.0239**
	(0.0335)	(0.0175)	(0.0401)	(0.00961)
Observations	915	915	915	915
R-squared	0.011	0.006	0.010	0.002

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

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

4.2.3 Being trained with a socially close peer is better.

We now use the randomness in the identity of the matched peer in order to understand the conditions under which pairing can be helpful. We first distinguish between being paired with someone who is a friend.

As seen from Table 4, being paired with a socially close individual is larger in magnitude than being paired with a non-friend for three crucial outcomes: readiness to invest, business index, and take-up index. We are not powered to show that these differences are statistically significant but the p-value of the difference for readiness to invest is 0.11. At the same time, the difference between being paired with a socially close peer and being trained alone is significant for readiness to invest (p-value=0.09).

Alternative definition of friendship: We also re-run the regression with a more conservative definition of social closeness by defining a friend as an individual with network distance equal to one. In this case, we find that those trained with friends are significantly more likely to be ready to invest compared to those trained with non-friends (p-value=0.06).

Table 4: Impact of being trained with a friend

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	-0.0554	0.0680	0.343***	0.288***
	(0.0784)	(0.0474)	(0.0891)	(0.0876)
Paired (Friend)	-0.0669	0.153***	0.401***	0.326***
	(0.0777)	(0.0352)	(0.0863)	(0.0834)
Paired (Not Friend)	-0.0187	0.0902**	0.311***	0.323***
	(0.0812)	(0.0396)	(0.0859)	(0.0737)
Constant	0	0.704***	0	0
	(0.0529)	(0.0283)	(0.0716)	(0.0659)
Observations	1,164	1,163	1,151	1,148
R-squared	0.001	0.017	0.036	0.035
Treatment 1== Paired with friend	0.905	0.0882	0.453	0.585
Treatment 1==Paired with nonfriend	0.713	0.579	0.708	0.565
Paired with friend==Paired with nonfriend	0.558	0.109	0.235	0.933

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

Notes: This regression treats the control group as the base category. Treated with a friend implies observations in Treatment 2 and Treatment 3 that are paired with someone with a social distance less than equal to two. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

4.2.4 Being trained with a socially close and central peer is better.

Finally, we investigate who benefits from the intervention by additionally considering the centrality of the individual one is matched with (i.e. the number of their network connections). As discussed earlier, this implies a two by two categorisation by social distance and centrality i.e. socially close peers that are more central than the individual, socially close peers that are less central, non-close individuals that are more central and non-close individuals that are less central.

We find that the training is more successful when individuals are paired with socially close individuals who are central in the social network. This is seen in Table 5 where we compare the pairs with those treated alone. "Friend X More Central" is an indicator variable that takes the value 1 if the pair is socially close and more central than i whereas "Friend X Less Central" is an indicator variable that takes the value 1 if the pair is socially close and less central than i. The effect of having a friend who is central in the pair positively affects readiness to invest, and take-up index when compared to Treatment 1– the difference is significant at 10% for readiness to invest and at 5% for the take-up index. The effect of being matched with such a person rather than being trained alone is higher in magnitude for the business index as well, albeit not statistically significant.

Table 5: Impact of training with different friend type

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Friend x More Central	-0.00185	0.114*	0.132	0.134**
	(0.133)	(0.0653)	(0.140)	(0.0636)
Friend x Less Central	-0.0133	0.0506	-0.0505	-0.0252
	(0.0835)	(0.0467)	(0.0751)	(0.0784)
Not Friend x More Central	-0.0211	0.0159	0.00700	0.0681
	(0.103)	(0.0570)	(0.138)	(0.0694)
Not Friend x Not Central	0.0634	0.0250	-0.0620	0.0258
	(0.125)	(0.0360)	(0.0778)	(0.0724)
Constant	0.0197	0.741***	0.189	0.255***
	(0.122)	(0.0612)	(0.114)	(0.0871)
Observations	681	682	671	670
R-squared	0.002	0.009	0.011	0.006
Friend x Less Central==Friend x More Central	0.904	0.258	0.240	0.0325
Friend x Less Central==Non Friend x Less Central	0.402	0.494	0.881	0.474
Friend x Less Central==Non Friend x More Central	0.940	0.518	0.673	0.177
More Central==Less Central	0.699	0.495	0.237	0.0614
Friend==Non Friend	0.727	0.132	0.408	0.849

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

Notes: This regression treats Treatment 1 as the base category. Friend includes pairs that have social distance less than or equal to two and nonfriend includes pairs with social distance greater than two. Friend x More Central equals 1 when the individual the respondent is matched with is a friend and more central than the respondent. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

4.2.5 Robustness

As shown in Table E.1 and E.2, the main results (studying the impact of the various treatment arms on endline outcomes) are robust to accounting for imbalance in baseline characteristics using the post double selection Lasso method in Belloni et al. (2014). Further, as shown in Table E.3 we find that the effect of being matched with a peer who is close and more central is significantly higher than Treatment 1 for take-up index (at 10%) after we account for any imbalance in baseline characteristics using the post double selection Lasso method in Belloni et al. (2014). Moreover, the effect is also significantly higher effect among those matched with a more central friend than a less central friend.

4.2.6 Alternative explanations for the friend-central effect

Before proceeding with the longer-term results, we also look at various characteristics beyond centrality to understand if there is any confounder that may be driving the effect of being paired with a central friend in the endline. This is to understand whether the effect of the measured network position of the peer persists even after controlling for baseline characteristics.

First, we 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 of these variables.[19] Table D.1 in the appendix shows these results where we compare Treatment 1 with various pair types and additionally control for similarity which is normalised to be between 0 and 1.[20] We find that the effect of being matched with a peer who is central on readiness to invest, business index, and take-up is still higher in magnitude than all other categories and is significant for the take-up index.

Next, we show in Table D.2 in the appendix that the effect magnitude of being matched with a central friend is higher than Treatment 1 (and significantly so for readiness to invest and the business index), even after controlling for whether the peer is in the upper caste, is unmarried, is older, has higher education, and has higher income. Importantly, this shows that the network characteristics of the peer matter even after controlling for other potential confounders.

4.3 Long Term Effects after 1 year of Training

In this section, we will be looking at outcomes based on a phone survey conducted in October 2023 i.e. one year after the training, with a random subset of the original sample comprising of ~ 750 individuals. We will compare the treatment group with those in the pure control villages. To ensure that the pure control villages are not systematically different from the treated villages, we show a balance test in Table B.3 where we find that almost all characteristics are balanced between individuals in these villages. Moreover, we show in Table B.2 that the different treatment groups are balanced in this follow-up sample along various baseline characteristics. As before, we will also show that results are robust to the inclusion of Lasso-selected controls following the procedure outlined in Belloni et al. (2014).

We will focus on the effect of the intervention on the treatment group. A brief discussion of the observed effects on the spillover group is presented in the appendix in Section F.

We find that about 3% of individuals in our follow-up sample opened up new businesses after one year of training. As shown in Table 6, we find that those who were treated alone or

[19]As before, we use the method proposed in Anderson (2008) to construct the index.
[20]This implies that similarity is equal to 1 for the most similar pairs in the sample.

in pairs were not any more likely to have opened up businesses when compared to those in pure control villages and, conditional on opening a new business, were not any more likely to have invested more in it. At the same time, we find that those who were treated in a pair have significantly higher self-reported agricultural profits than those who were treated alone – their agricultural profits are more than thrice as large as those treated alone.[21] Moreover, these individuals also have significantly higher agricultural investments when compared to the pure control group, as shown in the fourth column of Table 6.

The effect of being trained together on agricultural investments and profits 1 year later is consistent with the finding in the endline survey that those who were trained together were more likely to report wanting to open an agro-business and that individuals in these villages tend to have cooperative labour-sharing agreements with each other (Messerschmidt 1981, Sherpa 2005, Bhattarai 2006). We also do not detect any treatment effects on whether individuals maintain records of their agricultural expenses, making it unlikely that higher profits result from better record-keeping.

Table 6: Long-term effect on Business and Agriculture Outcomes.

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Treated alone (T1)	0.00591	8,815	1,093	21,594
	(0.0212)	(17,347)	(1,495)	(17,357)
Treated with a pair	-0.00955	26,669	99.44	28,855**
	(0.0185)	(17,452)	(1,232)	(14,083)
Constant	0.0395**	102,770***	2,034*	98,576***
	(0.0158)	(13,463)	(1,005)	(8,965)
Observations	750	734	750	589
R-squared	0.007	0.008	0.005	0.007
Paired=Nonpaired	0.407	0.0773	0.530	0.680

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. Agricultural profits, business investments, and agricultural investments are self reported and winsorised at 0.01% to exclude outliers. Investments in new businesses are non-zero only for those who opened a new business in the previous year. Standard errors are robust and clustered at the village level.

Table 7 shows that this translates into significantly higher monthly income for those who were treated in pairs, although the result is only significant at 10% and is not statistically distinguishable from the effect of being treated alone. Training also increases the amount of monthly savings, whether or not they have taken a loan, and willingness to take up a commitment savings account offered by the experimenter where the saved money could only be withdrawn for business purposes. These effects do not differ by whether the individual was trained alone or in a pair. However, those trained alone are significantly more likely to

[21] As shown in Table C.1 in the appendix, this effect is primarily driven by Treatment 3, which has significantly higher monthly profits compared to Treatment 1 but is not significantly different from Treatment 2.

have opened a new savings account when compared both with the control group and those who were treated in a pair.

Table 7: Long-term effect on Economic Outcomes.

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Treated alone (T1)	2,975	4,015***	0.107***	0.0874**	0.0633
	(2,561)	(1,280)	(0.0376)	(0.0381)	(0.0677)
Treated with a pair	4,436*	2,765*	0.0267	0.0759**	0.127**
	(2,372)	(1,373)	(0.0340)	(0.0322)	(0.0561)
Constant	22,921***	3,401***	0.197***	0.0674***	0.446***
	(1,837)	(825.5)	(0.0277)	(0.0186)	(0.0490)
Observations	749	744	752	752	751
R-squared	0.006	0.015	0.008	0.010	0.011
Paired=Nonpaired	0.344	0.400	0.0698	0.752	0.200

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. Monthly income and savings are winsorised at 0.01% to exclude outliers. Standard errors are robust and clustered at the village level.

4.3.1 Training improves long-term outcomes for those matched with socially close pairs

To explore the reasons behind the observed increase in monthly income, agricultural profits, and investments for those treated in a pair, we combine treatment arms 2 and 3 as before and first split them into two categories– whether pairs were matched to a friend or a non-friend. Then, we compare the treatment effects for those two categories with the treatment effects for those who were trained alone. Table 8 shows the results for business and agricultural outcomes and Table C.3 shows the results for other economic outcomes.

We find that agricultural profits are significantly higher for those who were paired with a friend compared to those who were treated alone. The treatment effect on those trained with a friend is about $\sim 50\%$ of the mean agricultural profits of those treated alone. While monthly income is higher in magnitude among those who were paired with a friend compared to those treated alone, we are not powered to detect differences. It is important to note that the significant average effect of the paired treatment on monthly income arises from those paired with a friend as the coefficient for those paired with a non-friend is negative.

4.3.2 Training improves long-term outcomes for those paired with less central individuals

Similarly, we split the paired treatment arms into two categories– paired with a more central friend or a friend with the same or lower centrality (in terms of number of connections, as before). These results are reported in Table 9 for business and agricultural outcomes and Table C.4 for other economic outcomes.

Table 8: Long-term effect on Business and Agriculture Outcomes by Friendship Status

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Paired (Friend)	-0.0105	38,393**	-1,184	3,539
	(0.0230)	(16,974)	(1,882)	(23,478)
Paired (Not Friend)	-0.0200	8,578	-1,030	8,165
	(0.0191)	(11,366)	(1,639)	(21,568)
Constant	0.0773***	76,564***	6,153***	100,021***
	(0.0205)	(22,753)	(2,010)	(25,469)
Observations	432	424	432	318
R-squared	0.006	0.020	0.008	0.003
Paired with friend==Paired with nonfriend	0.618	0.202	0.922	0.849

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. Agricultural profits, business investments, and agricultural investments are winsorised at 0.01% to exclude outliers. Investments in new businesses are non-zero only for those who opened a new business in the previous year. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

Unlike the immediate results in the endline survey, we find that being paired with a more central person is not helpful. In fact, we find that those who are treated with a less central person have significantly higher agricultural profits compared to those trained alone and those trained with a more central individual. This suggests that being matched with a more central individual may not be as beneficial in the long run but being matched with a friend is beneficial both in the short and long run. Moreover, these benefits of training do not translate into opening businesses but into investing more and earning more from agriculture. Even though it is not significant, the magnitude of agricultural investments for those matched with less central individuals is 14% higher than the agricultural investments of those matched with more central individuals.

These results are robust to the inclusion of Lasso-selected controls in accordance with the procedure outlined in Belloni et al. (2014). These results are shown in Tables E.4-E.9. Moreover, these results are also consistent with the endline specification shown in Tables C.5 and C.6. The tables show that, unlike other pair types, a non-central friend (i.e. Friend X Less Central) has significantly higher income and agricultural profits compared to those treated alone. These pair types are less likely to have opened a new business and earn more from agriculture instead.

Table 9: Long-term effect on Business and Agriculture Outcomes by Centrality of Peer

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Paired (More Central)	-0.0122	3,668	-2,198	-2,392
	(0.0226)	(10,986)	(1,817)	(23,923)
Paired (Less Central)	-0.0194	30,769***	-426.9	11,772
	(0.0205)	(10,459)	(1,792)	(18,221)
Constant	0.0731***	80,804***	6,744***	103,988***
	(0.0221)	(23,905)	(2,304)	(26,379)
Observations	436	428	436	322
R-squared	0.006	0.019	0.010	0.004
Paired with more central==Less central	0.745	0.0408	0.349	0.416

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. Agricultural profits, business investments, and agricultural investments are winsorised at 0.01% to exclude outliers. Investments in new businesses are non-zero for those who opened a new business in the last year. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

5 Mechanisms

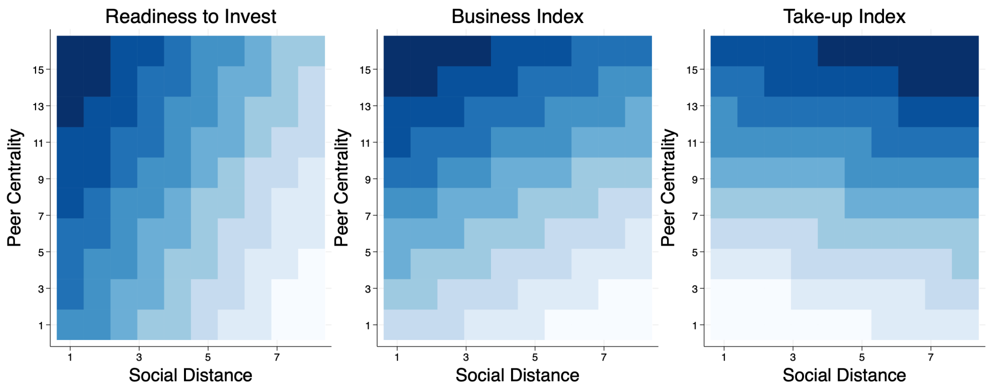
We find that being matched to a more central and socially close peer can help in the short term while being matched to a less central and socially close peer can help in the longer term. We will also reaffirm these results using alternative specifications (with higher statistical power) in Section 6 when we directly estimate peer effects and run dyadic regressions.

Our results so far are summarized in Figure 7, where we plot the predicted values of various short-term and long-term outcome variables across different levels of social distance and peer centrality, using a model trained on data from the field experiment. As shown in the figure, the effect of centrality reverses over time, with darker shades shifting from the top to the bottom. This indicates that while a more central peer is helpful in the short term, a less central peer is more beneficial in the long term.[22]

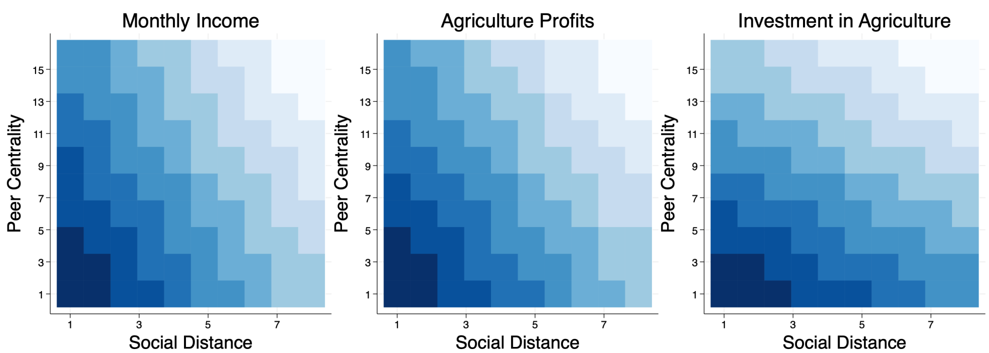
[22] Note that the effect of social distance on the take-up index is different from its effect on other outcomes. This is likely because receiving help from a socially close individual reduces the need to demand additional resources i.e. these two sources of support are substitutes.

Figure 7: Effect of Social Distance and Peer Centrality over Time

(a) Short Term



(b) Long Term



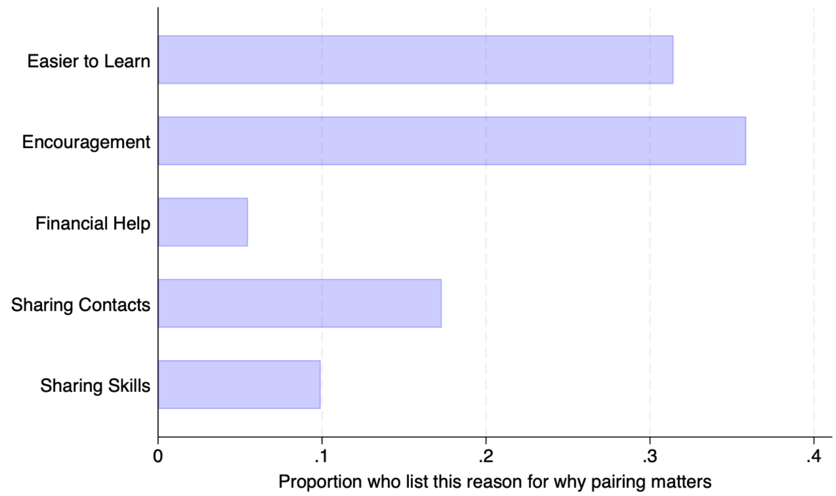
Notes: The above graph plots the predicted values of outcome variables for individuals in Treatments 2 and 3 from the endline and follow-up waves, computed at various values of social distance and peer centrality. Lighter colours correspond to lower predicted values of the corresponding variables. The model is estimated using the experimental data and is applied to a dataset where centrality and social distance are uniformly sampled between the minimum and maximum observed in the data.

In this section, we suggest that these effects arise due to motivation in the short term and ease of collaboration in the long run. We disentangle various possible channels through which these peers could be helpful to show this.

5.1 Participant Reports

First, we asked the perceptions of women who attended the training in pairs as to how pairing was useful rather than being training alone. Figure 8 shows the responses to this question. 33% of women paired report having received encouragement from their partnered peer. This is followed by 31% of the women who report that the training material was easier to grasp due to being paired in the training. About 20% say that the training helped because they shared network contacts. Importantly, the responses to this question do not differ by whether the individual was paired with a friend or with a more/less central individual than them. We now proceed with disentangling mechanisms that can explain how pairing can help and why the effects are heterogeneous by centrality and distance.

Figure 8: Why does pairing matter?



5.2 Do pairs learn better together?

5.2.1 Impact on knowledge after the training

The training could be helpful for peers as they can help each other learn the material better. To evaluate this mechanism, we first ask five short questions that measure knowledge gained in 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. [23] Column 1 of Table 10 shows that while the intervention improved the knowledge index, this does not differ by whether the individual was trained alone or with a peer.

5.2.2 Impact on the performance during the training

Next, we also measure knowledge gained by the training by two additional variables measured in exercises conducted during the training: game profit and yearly profit. Game profit records the performance on the first day of the training where women played an investment and saving game. Yearly Profit on the other hand was measured on the last day of the training,

[23] The questions are as follows: (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?, and (e) Above what break even percentage does the business become risky? The measure of knowledge is kept brief due to logistical constraints and is therefore likely to be noisy.

which is the amount of profit made in the exercise involving the creation of a business plan.

Columns 2 and 3 of Table 10 show that these endline outcomes do not differ by whether the individual was trained alone or with a peer. However, it is important to note that paired individuals make a profit equal to about double what is made by those treated alone, even though the difference is not statistically significant.

Table 10: Effect on Learning during and after the training.

VARIABLES	(1) Knowledge Index	(2) Profit (Game)	(3) Profit (Business Plan)
Treatment 1	0.955***		
	(0.130)		
Treated in a pair	0.925***	-6.208	465,002
	(0.135)	(83.79)	(459,637)
Constant	0	299.0***	458,013**
	(0.129)	(71.94)	(173,290)
Observations	1,204	725	723
R-squared	0.246	0.000	0.002
T1= Paired	0.608		

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

Notes: This regression treats the Control Group as the base category in Column 1 and Treatment 1 as the base category in Column 2. Standard errors are robust and clustered at the village level.

5.2.3 Impact on knowledge by type of peer

It could also be the case that women who attended the training with a friend may have been able to better discuss the material being taught. If this is the case then women treated with a friend should perform better in the knowledge and performance in the training. Table D.3 shows that the effect of being trained with a friend on knowledge during training, measured in terms of performance in the profit-making exercises, is much higher than treatment 1 and much higher than being trained with a non-friend, even though we are not powered to detect differences. Table D.4 shows that the effect is not very different by the type of peer in terms of a combination of social distance and centrality, even though the profit made in the business plan is almost 1.5-2 times higher among those matched with a central friend than those matched with other friend types.

5.3 Do pairs provide indirect value in terms of access to network connections?

Gaining access to network connections by pooling network contacts is the second channel that we look at. Treatment 3 i.e. where women were paired during the training and a connection module was designed precisely to check if this mechanism is at work. As previously discussed, this treatment fares better in terms of magnitude but the effects on most outcomes are not significantly different when compared to Treatment 1 or 2.

5.3.1 How do the number of pooled contacts depend on demographic and network characteristics?

On average, women in Treatment 3 pool 6.5 contacts. We find that there is no difference in the number of contacts pooled as a function of age, degree, education and income of the peer, as shown in Table D.5. However, being in the same caste has a significantly higher effect and leads to one additional friend being listed. Moreover, we find that pairs that have a larger gap in their number of network connections are significantly more likely to pool more contacts.

Finally, we do not find heterogeneous effects by the number of contacts pooled in Treatment 3 as shown in Table D.6. This suggests that pooling network contacts might not be a mechanism of interest.

5.4 Do pairs help each other share risk?

Pairing can also improve outcomes if the peer is less risk-averse or if the peer provides financial assistance. We do not find this to be the case and show it in three ways.

5.4.1 Heterogeneity by Risk Aversion

First, we do not find any heterogeneous effects by baseline risk aversion of the matched peer. As seen in Table D.7, being paired with a peer who is less risk-averse does not impact the outcomes of the training.

5.4.2 Impact on Network-Based Methods of Savings

Next, 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 do not find any effects on these variables. This suggests that the training, and especially the paired treatments, did not lead to an increased involvement with the network in terms of saving together and sharing risks. These results are shown in Table 11.

Table 11: Treatment Effects on Network-Based Measures of Saving

VARIABLES	(1) Number of Savings Groups	(2) Joined Cooperative in last year
Spillover	-0.221	0.0558
	(0.257)	(0.0559)
Treated alone (T1)	0.168	0.0464
	(0.268)	(0.0350)
Treated in a pair	0.0609	0.0178
	(0.244)	(0.0346)
Constant	1.916***	0.0955***
	(0.229)	(0.0270)
Observations	751	752
R-squared	0.009	0.004
Spillover=T1	0.0112	0.868
T1= Paired	0.273	0.312
Spillover=Paired	0.0123	0.521

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

Notes: This regression treats the Control Group as the base category. Standard errors are robust and clustered at the village level.

5.4.3 Alternative Measures of Centrality

Finally, we also do not find effects of a peer being more central when we use eigenvector centrality as a measure of influence. This also suggests that risk-sharing might not be a relevant mechanism. As seen in Table D.8, when we define friend central differently using eigenvector centrality instead of degree centrality, it does not have differential effects compared to the other categories or compared to being treated alone. The effect of friend central vanishes highlighting that the number of connections of the peer matters as opposed to their eigen-vector centrality.

5.5 Do pairs encourage each other?

Another possible channel driving peer effects in the endline could be encouragement. This effect is in line with the fact that friends who have more friends are likely to be 'popular' and, therefore, are better at encouraging. The majority of individuals say that pairing benefited them because of encouragement.

First, we find that those with lower income aspirations in the baseline are significantly more likely to say that they were encouraged by their peer. Next, we find that those matched with a peer with a higher indegree (i.e. those who are listed as a friend by others and are 'popular') had higher treatment effects on endline outcomes as shown in Table 12. These heterogeneous effects do not exist when we consider the peer's outdegree (i.e. the number of friends the matched peer listed). This suggests that being matched with a person perceived to be popular by others is helpful in improving endline outcomes.

Table 12: Treatment effect by Peer's Indegree and Outdegree

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Heterogeneity by Peer's Indegree				
Treatment 1	-0.0975	0.0680	0.343***	0.288***
	(0.0970)	(0.0474)	(0.0891)	(0.0876)
Treatment 2 and 3	-0.0219	0.0901**	0.280***	0.237***
	(0.0864)	(0.0339)	(0.0879)	(0.0838)
T2/T3 X Peer's Indegree	-0.00362	0.00830**	0.0223*	0.0260**
	(0.0131)	(0.00403)	(0.0116)	(0.0117)
Constant	0	0.704***	0	0
	(0.0578)	(0.0283)	(0.0716)	(0.0659)
Observations	1,175	1,170	1,157	1,155
R-squared	0.001	0.015	0.036	0.036
Heterogeneity by Peer's Outdegree				
Treatment 1	-0.0975	0.0680	0.343***	0.288***
	(0.0970)	(0.0474)	(0.0891)	(0.0876)
Treatment 2 and 3	0.161	0.129***	0.350***	0.384***
	(0.145)	(0.0444)	(0.109)	(0.0899)
T2/T3 X Peer's Outdegree	-0.0632	-0.00494	-0.00163	-0.0231
	(0.0468)	(0.0123)	(0.0286)	(0.0201)
Constant	0	0.704***	0	0
	(0.0578)	(0.0283)	(0.0716)	(0.0659)
Observations	1,175	1,170	1,157	1,155
R-squared	0.003	0.014	0.034	0.034

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

Notes: This regression treats the Control Group as the base category. Standard errors are robust and clustered at the village level.

Finally, we also find that being paired translates into higher measures of self-efficacy for those who were paired, as shown in Table 13. In particular, we find that those matched with a peer are significantly more likely to believe that they are confident to deal with adverse events and can accomplish their goals when compared to the control group (at 10%). The effect, however, is not significantly different from those who were trained alone.

5.6 Are pairs likely to help each other and/or collaborate in future?

Lastly, we find that individuals paired with friends report being more likely to want to open a business together during the endline when compared to those paired with non-friends. As shown in Table 14, we find that individuals are significantly more likely to report wanting to meet with their matched peer in future if they are a friend rather than a non-friend. Moreover, when asked to report if they were likely to open a business with the individual

Table 13: Treatment effects on Self Efficacy

VARIABLES	(1) Can Manage Difficult Problems	(2) Feels Confident to deal with events	(3) Can Accomplish Goals
Treatment 1	0.0406	0.0623	0.0635
	(0.0632)	(0.0486)	(0.0533)
Treated in a pair	0.0426	0.0719*	0.0824*
	(0.0402)	(0.0394)	(0.0415)
Constant	0.308***	0.285***	0.289***
	(0.0346)	(0.0316)	(0.0349)
Observations	1,202	1,203	1,202
R-squared	0.002	0.005	0.006
T1= Paired	0.977	0.839	0.710

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

Notes: This regression treats the Control Group as the base category. Standard errors are robust and clustered at the village level.

they were paired with, being friends has a strong and significant effect. Both these effects are significantly higher for friends than non-friends. This reinforces the argument that distance between pairs matters as they make plans together. Centrality does not have a differential effect on these outcomes.

Table 14: Effect of Pair type on Wanting to start a business together.

VARIABLES	(1) Pairs will meet in Future	(2) Pairs will start Business Together
Friend x More Central	0.113**	0.143**
	(0.0444)	(0.0655)
Friend x Less Central	0.105***	0.155**
	(0.0357)	(0.0559)
Not Friend x More Central	0.0122	0.0742
	(0.0268)	(0.0563)
Degree	0.00496	0.0116
	(0.00427)	(0.0131)
Constant	0.813***	0.229**
	(0.0461)	(0.0942)
Observations	436	435
R-squared	0.029	0.022
More Central==Less Central	0.671	0.470
Friend==Non Friend	0.0107	0.0320

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

Notes: This regression treats Treatment 1 as the base category. Both the variables are self-reported willingness to meet and start a business together respectively, as measured during the endline survey. Standard errors are robust and clustered at the village level.

One year later, we find that individuals who are trained are also significantly more likely to report discussing business-related concerns with their social networks. This is shown in Table

15. When asked whether they spoke to their matched peer, we find that 35% of individuals in treatment 2 and 41% of individuals in treatment 3 report speaking to their matched peer. When asked what they spoke about, 85% report speaking to the peer for advice, 10% report interacting to borrow/lend money, and 5% report speaking about setting up businesses.

Table 15: Follow-up survey and Network Communication

VARIABLES	(1) Talk to anyone about business	(2) Talk to matched peer
Treated alone (T1)	0.0548*	
	(0.0280)	
Treated in a pair (T2/T3)	0.0874***	
	(0.0297)	
Treatment 3		0.0607
		(0.0596)
Constant	0.0226	0.353***
	(0.0154)	(0.0527)
Observations	751	260
R-squared	0.019	0.004
Paired=Nonpaired	0.325	

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. Both the variables are self-reported and show whether an individual talked to anyone or the matched peer about businesses during the one year after the intervention. Standard errors are robust and clustered at the village level.

Importantly, as shown in Table 16, among those who did speak to their matched peer, those matched with a weakly more central peer were less likely to talk about borrowing or businesses and more likely to take general advice. This effect is not statistically significant. However, it is similar to the significant effect of same-caste peers who are also more likely to talk than those paired with an individual of a different caste. This highlights the potential role played by the ease of collaboration among individuals of similar social standing and corroborates the previous empirical findings.

6 Peer Effects and Policy Counterfactuals

6.1 Framework

We now present a simple framework in which agents decide how much entrepreneurial effort to exert depending on their own characteristics and the effort exerted by their matched peer. We specifically focus on the role of heterogeneity in the network characteristics of the matched peer on the individual's choices and outcomes. There are two variables that capture this heterogeneity: the distance between the participant and the person they are matched with and the network centrality of the person they are paired with.

Table 16: Talking to Matched Peer

VARIABLES	(1) Talk to matched peer	(2) Talk to matched peer	(3) Talk to matched peer for borrowing or business	(4) Talk to matched peer for borrowing or business
Paired (More Central)	0.0908		-0.0891	
	(0.0542)		(0.0562)	
Same Caste		0.117*		0.0960
		(0.0616)		(0.0642)
Constant	0.345***	0.291***	0.180***	0.0625
	(0.0447)	(0.0592)	(0.0605)	(0.0514)
Observations	246	256	94	98
R-squared	0.008	0.010	0.017	0.010

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

Notes: This regression treats those matched with a less central peer in Treatments 2 and 3 as the base category. Both the variables are self-reported and show whether an individual talked to the matched peer about businesses during the one year after the intervention and if so, whether they spoke about business-related matters. Standard errors are robust and clustered at the village level.

Consider the following utility function where agent i chooses the level of effort depending on their own characteristics, peer's characteristics, and peer's effort choice.

$$
\begin{aligned}
U(e_i) = \quad & \theta e_i + \underbrace{\theta_{\mathbf{o}}\ \mathbf{X_i}\ e_i}_{\text{returns to effort depending on own characteristics}} + \underbrace{\theta_{\mathbf{p}}\ \mathbf{X_j}\ e_i}_{\text{returns to effort depending on peer characteristics}} \\
& + \underbrace{\theta_e\ e_i e_j}_{\text{returns depending on peer effort}} + \underbrace{\theta_d\ d_{ij}\ e_i e_j}_{\text{additional effect of social distance}} \\
& + \underbrace{\theta_c\ \phi_j\ e_i e_j}_{\text{additional effect of peer centrality}} - \underbrace{c(e_i)}_{\text{cost of effort}}
\end{aligned}
$$

where $\mathbf{X_i}$ and $\mathbf{X_j}$ are vectors of individual and peer characteristics, respectively, d_{ij} is the distance between i and j, ϕ_j is the number of network connections of the agent j, and the effort cost is convex, that is, $c''(e_i) > 0$. We also assume that $\theta_e>0$ so that agents are likely to exert more effort if their matched peer does so. Finally, θ_d and θ_c can be positive or negative, and importantly, the sign can differ over time. We do not add a time-subscript here but it is important to note that these complementarities could change over time. For example, in line with our empirical findings, we expect θ_c to be positive in the short term and negative or zero in the long term. The expression for optimal effort e_i^* can be written as follows:

$$c'(e_i^*) = \theta + \theta_{\mathbf{o}}\mathbf{X_i} + \theta_{\mathbf{p}}\mathbf{X_j} + \big(\theta_e + \theta_d d_{ij} + \theta_c \phi_j\big)e_j^*$$

Note that social distance and peer centrality affect the strength of the complementarity in effort choices, thereby affecting the agent's choice of effort. For instance, if $\theta_e > 0$, $\theta_d < 0$, and $\theta_c > 0$, then an individual exerts higher effort if their peer does so, and more so if the peer is socially close and more central in the network. Alternatively, if $\theta_e > 0$, $\theta_d < 0$, and $\theta_c < 0$, then an individual exerts higher effort if their peer does so, and more so if the peer is socially close but less central in the network. If $\theta_e > 0$, $\theta_d = 0$, and $\theta_c = 0$, then the strength of effort complementarity does not depend on social distance and network centrality.

6.1.1 Correlational Evidence

We estimate the above first-order condition to assess if the magnitude of strategic complementarities might differ based on the network positions of the paired individuals. It is important to note that while random pairing allows us to address concerns about correlated effects observed in Manski (1993), the lack of exogenous variation in effort does not allow us to solve the reflection problem highlighted in Manski (1993). However, we first conduct this exercise to check if the correlations are in line with the predictions in the preceding empirical analysis. Then, we address this concern by testing how endline outcomes of peers affect the respondent's follow-up outcomes, thereby addressing the reflection problem. Moreover, in the next subsection, we will estimate dyadic regressions that will make use of the randomisation in our experiment and allow us to run policy counterfactuals.

We first estimate the equation for the first-order conditions for the endline and follow-up waves separately because of the empirical analysis that has previously shown that centrality and distance play a different role depending on the time frame when effects are observed. We use data from the paired treatments (i.e. Treatments 2 and 3) for this exercise. We construct a measure of effort to capture short-term effects by constructing a standardised Anderson Index (Anderson 2008) combining all the endline variables that enter the business, take-up, and knowledge index, along with readiness to invest. For the follow-up, we combine several variables including business investments, agricultural investments, agricultural profits, savings amount, loan amount, income aspirations, and whether individuals have spoken to anyone about opening a business.[24]

[24]We shift the indices by the minimum value after standardisation i.e. generate $z_new= z + z_{min}$ to ensure that own outcomes and peer outcomes are always positive. This will allow us to interpret the coefficients on the distance and centrality interaction terms.

Using the indices, we then estimate the following equation that follows from the derived first order condition once we assume that the cost function is quadratic and $c(e_i) = \frac{1}{2}e_i^2$.

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

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.

Table 17: Effect of Distance and Centrality in the Short and Long Run

	Short Run		Long Run	
	(1)	(2)	(3)	(4)
VARIABLES	Effort	Effort	Effort	Effort
θ_e: Peer Effort	0.102	0.0592	0.251***	0.204
	(0.0769)	(0.0772)	(0.0628)	(0.122)
θ_d: Distance X Peer Effort	-0.000460	0.00278	-0.0624***	-0.0619**
	(0.00518)	(0.00468)	(0.0196)	(0.0234)
θ_c: Peer Degree X Peer Effort	0.00610*	0.00608	0.00694	0.0153
	(0.00325)	(0.00366)	(0.0261)	(0.0185)
Constant	2.580***	2.906***	1.532***	2.367**
	(0.214)	(0.403)	(0.255)	(0.895)
Observations	440	440	180	180
R-squared	0.022	0.129	0.042	0.202

Notes: The above table presents the results when we estimate the model predicted optimal effort equation. The first two columns use data from the endline survey while the last two columns use data from the follow-up survey. Columns (2) and (4) additionally control for individual and peer characteristics. The sample for the paired treatments in the follow-up is smaller here as we only include cases where we observe both the individual and their peer after 1 year. We only use data from Treatments 2 and 3 since these are the only treatments where pairs are observed.

The results are shown in Table 17. We find suggestive evidence of heterogeneity in the magnitude of strategic complementarity depending on how close and central the peers are. Moreover, we find that this heterogeneity varies over time. As shown in the table, we find that the average outcomes are higher when agents are connected to more central peers in the endline but to closer peers in the follow-up. The effect of being paired with central peers disappears in the long run as social distance becomes more relevant.

6.1.2 Identification of Peer Effects

Next, as shown in Table 18, we identify peer effects by regressing the follow-up outcomes of the individual on the endline effort of their peers. This addresses the reflection problem Manski (1993), which is otherwise not addressed solely by random matching.

As shown in the table, we continue to find that social distance reduces the effect of the peer on follow-up outcomes. This effect persists even after we control for the individual's and peers' demographic characteristics and the individual's own outcome in the endline. This is shown in Column 2.[25] Moreover, the coefficients on both interactions (with network centrality and distance) are always negative, consistent with the heterogeneity of peer effects discussed earlier.

Table 18: Identifying Peer Effects by regressing Follow-up Outcome on Peer Outcome in the Endline.

VARIABLES	(1) Outcome (Follow-up)	(2) Outcome (Follow-up)
Peer Outcome (Endline)	0.263**	0.123
	(0.122)	(0.141)
Distance X Peer Outcome	-0.0454**	-0.0420**
	(0.0165)	(0.0199)
Peer Degree X Peer Outcome	-0.00811	-0.00176
	(0.00922)	(0.00649)
Own Outcome (Endline)		0.175
		(0.173)
Constant	1.483***	1.977**
	(0.334)	(0.905)
Own Characteristics	No	Yes
Peer Characteristics	No	Yes
Observations	178	178
R-squared	0.034	0.198

6.2 Counterfactuals: Network-Based vs Random Pairing

What would happen if we paired individuals with a central person or with a friend? How would this compare to pairing individuals randomly, i.e. without leveraging the network structure? In order to capture the magnitude of the effect of a policy that achieves this, we first run a dyadic regression where we exploit the random variation in distance and centrality among the dyads. The results of the dyad level regressions are presented in Table

[25]We estimate peer effects using these outcomes since we detect positive treatment effects on them in the follow-up survey and wish to understand how peer characteristics affect this.

19 where, consistent with our empirical results, we find that dyads that have a larger gap in their centrality perform significantly better in the short run while dyads that socially are close perform significantly better in the long run. The outcome variable used here is the standardised index of the endline and follow-up covariates as specified in the previous section. We now proceed to illustrate two simple counterfactual policies that compare network-based pairing with random pairing.

Table 19: Effect of Distance and Centrality on Dyad-Level Outcomes

VARIABLES	(1) Dyad Outcome (Short Run)	(2) Dyad Outcome (Long Run)
Close (Distance<=2)	0.0490	0.451*
	(0.0510)	(0.244)
Centrality Gap	0.00150**	-0.00358
	(0.000715)	(0.00795)
Constant	2.937***	1.555***
	(0.0397)	(0.179)
Observations	220	90
R-squared	0.006	0.046

Notes: The above table presents the results of a dyadic regression where we regress the average value of the outcome for the dyad on dyad-specific network variables. The centrality gap is equal to the quadratic difference between the network degrees of the individual and the person they are matched with. The first column uses data from the endline survey while the second column uses data from the follow-up survey. We only use data from Treatments 2 and 3 since these are the only treatments where pairs are observed. The sample for the paired treatments in the follow-up is smaller here as we only include cases where we observe both the individual and their peer after 1 year.

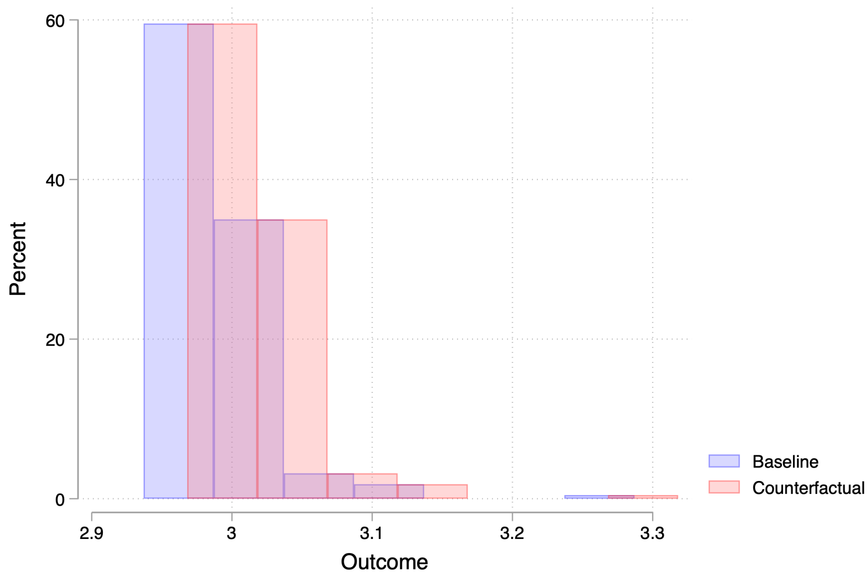
6.2.1 Policy Counterfactual 1: Pairing with central individuals vs Random Pairing

We use the endline data to show that under the counterfactual policy where individuals were always trained with someone such that the gap between their centralities is one standard deviation larger, we would increase the standardised endline outcomes index by 0.8 standard deviations compared to the baseline case where they are treated with randomly chosen individuals. Figure 9 plots the histogram of this index for the baseline and counterfactual cases respectively. As shown in the figure, we find that pairing with more central individuals shifts the distribution of the desired outcomes rightwards.

6.2.2 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 feasible pairings remain unclear. How many such pairs can we form within a village, given the fixed social network?

Figure 9: Counterfactual 1: Matching with Central Individuals vs Random Matching



We simulate 10,000 counterfactual reassignments of peer pairings within each village and find that strategic pairings based on centralities are indeed feasible and can generate high returns. For each village, we first grouped individuals into randomly chosen dyads, computed the gaps in centralities of each dyad, predicted dyad outcomes in accordance with the coefficient on centrality gaps obtained above, and then repeated this process 10,000 times.

We find that, on average, the best possible average endline outcome in a village across all counterfactual reassignments (i.e. the average outcome with the highest average centrality gap between the pair) is 8 times the worst possible outcome (i.e. with the lowest average centrality gap). This ratio is plotted in Figure 10 for all villages in our data.

This confirms that it is indeed possible to strategically pair individuals within villages using their centralities in order to maximise treatment effects.

6.2.3 Policy Counterfactual 2: Pairing with friends vs Random Pairing

Next, we use the follow-up data to show that under the counterfactual policy where individuals were always trained with their friends, we would increase average outcomes by 1.07 standard deviations compared to the baseline case in which they are treated with randomly chosen individuals. Figure 11 plots the histogram of the outcomes for the baseline and counterfactual cases respectively. As shown in the figure, we find that matching with socially close individuals shifts the distribution of the desired outcomes rightwards.

Figure 10: Effect of Counterfactual Pair Reassignments on Endline Outcomes

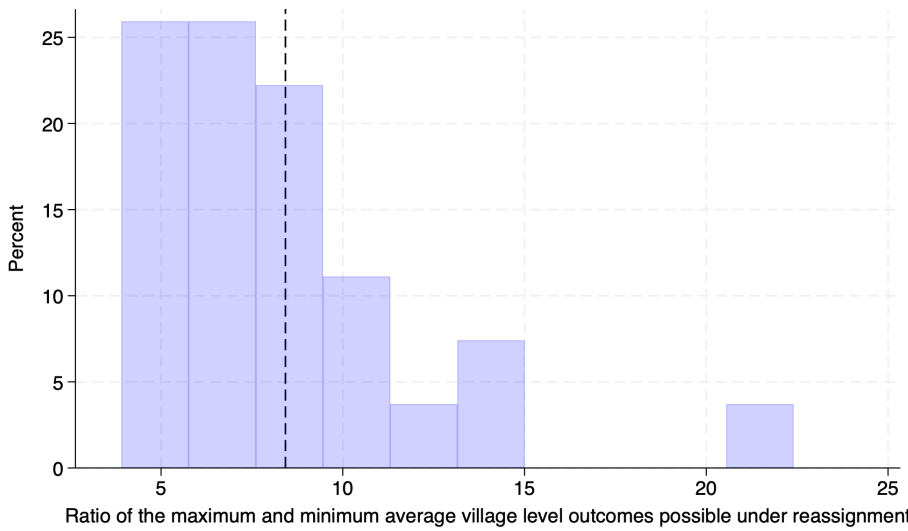
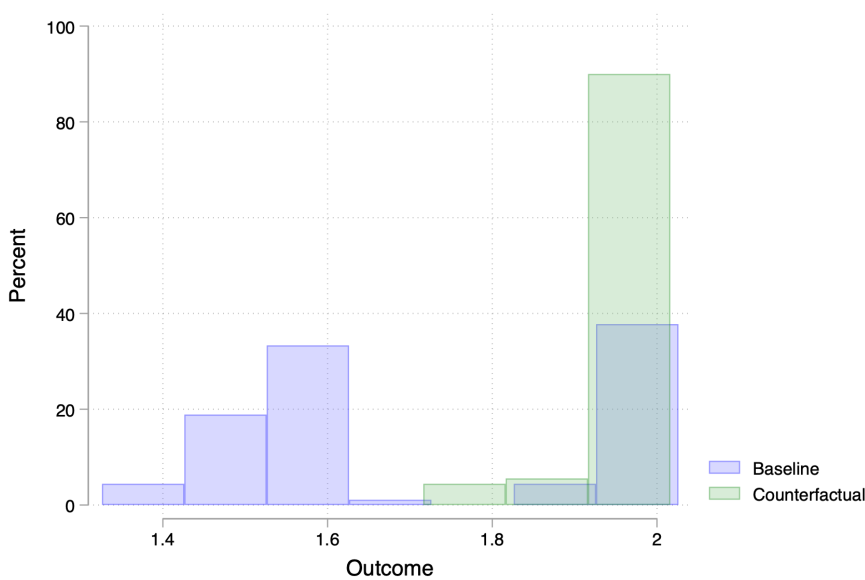


Figure 11: Counterfactual 2: Matching with Friends vs Random Matching



These counterfactuals suggest that there is value in pairing women in such training programs using information about their network positions. Moreover, the primary aim of the training might determine which of the two aspects i.e. closeness or centrality to focus on. For instance, if the aim is to improve short-term motivation, then popular individuals can help. However, in order to improve long-term outcomes, it might be better to match with socially close and less central individuals.

7 Conclusion

We implement a randomized controlled trial to show that strategic networking is beneficial compared to having interactions at random. We show that the networking strategy must evolve over time and its benefits are primarily direct as opposed to using networking to harness a wider social network. We exploit the experiment to identify peer effects and reveal substantial heterogeneity in terms of network closeness and connectedness. Finally, we run counterfactual exercises to show that such strategic networking is not just feasible in our context but can also generate high returns i.e. peer effects can be strategically leveraged to help women succeed as entrepreneurs. However, the manner in which networks can be used to improve outcomes must depend on the underlying objective eg: whether it is to boost short-term motivation or improve the probability of long-term collaboration.

Our field experiment is designed to disentangle the effect of having different types of peers from the network. We show that networking with specific kinds of pairs (in terms of their network centrality and social distance) can help improve outcomes relative to being trained alone. In particular, we find that being paired with a friend i.e. someone with low social distance, helps improve outcomes in the short run and long run. On the other hand, being paired with a more central person i.e. someone who is more popular in the network, helps improve outcomes in the short run but does not benefit individuals in the long run. We suggest that central friends can provide short-term motivation but non-central friends are easier to collaborate with in the longer term.

We also use our experiment to compare the direct value of peers to their indirect value in terms of providing access to the wider social network. To do this, we study the effect of a treatment arm that makes the peer's social network more salient and encourages the pair to share their network contacts. While contact pooling has a higher effect on certain outcomes, this additional effect is not significant when compared to the arm where individuals are paired without pooling contacts. This suggests that the direct value of being connected with a friend can be more important than the indirect in terms of increasing the program's effectiveness. This opens the door for future interventions to test this mechanism and assess if it can lead to the formation of useful network ties.

Appendix

A Baseline Tables

Table A.1: Summary Statistics

	Mean	SD
Age	37.98	(10.85)
Marital Status		
Divorced	0.00141	(0.0375)
Married	0.918	(0.274)
Unmarried	0.0669	(0.250)
Widow	0.0134	(0.115)
Education		
No Education	0.326	(0.469)
Informal Education	0.126	(0.332)
Primary (Class 1-5)	0.155	(0.362)
Secondary (Class 6-10)	0.249	(0.432)
Higher Education (Class 11, 12)	0.104	(0.305)
University	0.0402	(0.196)
Belongs to Upper Caste	0.334	(0.472)
Degree Centrality	4.604	(2.167)
Own Non Agr. Business (Yes/No)	0.220	(0.415)
Reasons to not Open Business		
Have no skills	0.496	(0.500)
Feel not Capable	0.277	(0.448)
Financial Reasons	0.239	(0.427)
No support from family	0.0129	(0.113)
Willingness to Open Non Agr. Business	0.419	(0.494)
Risk Aversion (1-6)	4.610	(1.406)
Aspirations- Annual Non Agr. Investments	592607.8	(2991506.9)
Aspirations- Monthly Income	556120.4	(13902715.0)
Aspirations- Annual Agr. Investments	307888.2	(1716845.5)
Income Aspirations > Current Income	0.845	(0.362)
Non Agr Exp Asp > Current Exp	0.246	(0.431)
Observations	2840	

Notes: The above table reports summary statistics (i.e. mean and standard deviation) for various baseline characteristics. Degree Centrality measures the number of network connections in the village social network.

Table A.2: Correlations between number of network connections and baseline characteristics.

	Degree Centrality
Age	-0.00765
Marital Status	
Divorced	0.0460*
Married	0.0823***
Unmarried	-0.0857***
Widow	-0.0250
Education	
No Education	-0.0731***
Informal Education	0.0439*
Primary (Class 1-5)	0.0189
Secondary (Class 6-10)	0.0858***
Higher Education (Class 11, 12)	-0.0422*
University	-0.0588**
Belongs to Upper Caste	0.0700***
Own Non Agr. Business (Yes/No)	-0.0563**
Willingness to Open Non Agr. Business	0.0157
Risk Aversion (1-6)	-0.0690***
Aspirations- Annual Non Agr. Investments	-0.0322
Aspirations- Monthly Income	-0.0255
Aspirations- Annual Agr. Investments	0.00108
Income Aspirations > Current Income	-0.0163
Non Agr Exp Asp > Current Exp	-0.0290

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

Table A.3: Correlations between whether or not an individual has opened a business and their baseline characteristics.

	Own Non Agr. Business
Age	-0.0944***
Marital Status	
Divorced	0.00270
Married	-0.0354
Unmarried	0.0499**
Widow	-0.0250
Education	
No Education	-0.181***
Informal Education	-0.00791
Primary (Class 1-5)	0.00828
Secondary (Class 6-10)	0.00354
Higher Education (Class 11, 12)	0.129***
University	0.221***
Belongs to Upper Caste	-0.00835
Degree Centrality	-0.0563**
Eigen Vector Centrality	-0.0569**
Risk Aversion (1-6)	-0.0910***
Aspirations- Annual Non Agr. Investments	0.138***
Aspirations- Monthly Income	0.0127
Aspirations- Annual Agr. Investments	-0.0244
Income Aspirations > Current Income	-0.0303
Non Agr Exp Asp > Current Exp	0.374***

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

Table A.4: Correlations between whether or not an individual is willing to open a business and their baseline characteristics.

	Willingness to Open Non Agr. Business
Age	-0.391***
Marital Status	
Divorced	0.0187
Married	-0.0508*
Unmarried	0.0965***
Widow	-0.0826***
Education	
No Education	-0.300***
Informal Education	-0.104***
Primary (Class 1-5)	0.0367
Secondary (Class 6-10)	0.251***
Higher Education (Class 11, 12)	0.167***
University	0.100***
Belongs to Upper Caste	-0.0558**
Degree Centrality	0.0157
Eigen Vector Centrality	0.00563
Risk Aversion (1-6)	-0.103***
Aspirations- Annual Non Agr. Investments	0.104***
Aspirations- Monthly Income	-0.0188
Aspirations- Annual Agr. Investments	0.0428*
Income Aspirations > Current Income	0.0129
Non Agr Exp Asp > Current Exp	0.212***

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

B Balance and Attrition

Table B.1: Balance Test for the Endline Sample

	(0) Control	(1) T1	(2) T2	(3) T3	(0) vs. (1), p-value	(0) vs. (2), p-value	(0) vs. (3), p-value	(1) vs. (2), p-value	(1) vs. (3), p-value	(2) vs. (3), p-value
Income	23199.708	26253.382	24382.716	28570.368	0.299	0.573	0.065	0.488	0.457	0.167
Income Source-Agri.	0.912	0.847	0.823	0.850	0.061	0.136	0.104	0.684	0.922	0.641
Income Source-Business	0.102	0.069	0.078	0.107	0.346	0.495	0.920	0.761	0.165	0.336
Income Source-Job	0.029	0.008	0.029	0.034	0.003	0.973	0.823	0.139	0.247	0.819
Income Source-Remit.	0.015	0.020	0.008	0.000	0.643	0.188	0.021	0.286	0.062	0.151
Income Source-Other	0.052	0.044	0.033	0.043	0.663	0.241	0.625	0.481	0.939	0.571
Age	40.040	39.496	37.613	38.940	0.653	0.036	0.347	0.055	0.537	0.100
Elementary Education	0.165	0.173	0.177	0.154	0.749	0.631	0.692	0.921	0.539	0.493
Higher Education	0.079	0.077	0.091	0.094	0.906	0.680	0.521	0.598	0.299	0.889
Informal Education	0.319	0.282	0.292	0.291	0.451	0.399	0.470	0.838	0.879	0.964
Univeristy Education	0.006	0.004	0.021	0.009	0.720	0.079	0.767	0.076	0.554	0.273
Secondary Education	0.242	0.270	0.267	0.291	0.404	0.508	0.175	0.946	0.527	0.527
Degree	4.502	5.373	5.264	5.202	0.001	0.002	0.002	0.689	0.409	0.793
Brahmin	0.061	0.085	0.086	0.107	0.560	0.491	0.303	0.921	0.390	0.211
Chhetri	0.228	0.323	0.284	0.321	0.306	0.528	0.339	0.277	0.962	0.334
Newar	0.443	0.371	0.403	0.393	0.574	0.764	0.711	0.328	0.561	0.678
Tamang	0.205	0.145	0.144	0.150	0.408	0.347	0.385	0.979	0.910	0.876
Dalit	0.048	0.028	0.012	0.009	0.467	0.172	0.157	0.404	0.382	0.687
Other	0.017	0.048	0.070	0.021	0.137	0.046	0.645	0.530	0.240	0.114

Notes: The balance tests compare characteristics of the control group with Treatment 1, Treatment 2, and Treatment 3 for the endline sample. The p-values indicate if the difference is statistically significant.

Table B.2: Balance Test for the Follow-up Sample

	(1) Pure Control	(2) Spillover	(3) T1	(4) T2/T3	(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
Income	24220.194	27369.437	28594.149	26531.553	0.472	0.288	0.414	0.806	0.830	0.551
Income Source-Agri.	0.921	0.899	0.845	0.827	0.643	0.142	0.068	0.255	0.237	0.742
Income Source-Business	0.079	0.067	0.045	0.037	0.750	0.319	0.145	0.425	0.208	0.595
Income Source-Job	0.036	0.050	0.013	0.033	0.426	0.026	0.842	0.045	0.464	0.225
Income Source-Remit.	0.006	0.025	0.032	0.003	0.219	0.135	0.667	0.739	0.107	0.102
Income Source-Other	0.042	0.059	0.026	0.033	0.418	0.412	0.576	0.154	0.162	0.674
Age	39.861	41.403	38.897	38.040	0.483	0.611	0.323	0.115	0.039	0.276
Elementary Education	0.230	0.101	0.206	0.177	0.015	0.598	0.189	0.026	0.042	0.462
Higher Education	0.079	0.101	0.058	0.110	0.646	0.664	0.494	0.217	0.756	0.030
Informal Education	0.315	0.345	0.258	0.267	0.709	0.450	0.502	0.157	0.129	0.874
Univeristy Education	0.006	0.000	0.006	0.013	0.307	0.965	0.377	0.334	0.028	0.464
Secondary Education	0.261	0.210	0.310	0.300	0.316	0.347	0.413	0.036	0.038	0.840
Degree	4.698	4.376	5.614	5.196	0.182	0.001	0.022	0.000	0.000	0.111
Brahmin	0.073	0.092	0.090	0.097	0.828	0.813	0.756	0.959	0.900	0.796
Chhetri	0.212	0.252	0.348	0.313	0.807	0.388	0.538	0.130	0.257	0.402
Newar	0.418	0.454	0.342	0.413	0.872	0.709	0.982	0.174	0.547	0.156
Tamang	0.218	0.151	0.148	0.133	0.541	0.503	0.400	0.949	0.668	0.544
Dalit	0.067	0.025	0.019	0.007	0.296	0.298	0.171	0.647	0.358	0.407
Other	0.012	0.025	0.052	0.037	0.516	0.194	0.134	0.209	0.283	0.515

Notes: The balance tests compare characteristics of those in the pure control group with those in the spillover group, Treatment 1, and paired treatment arms for the follow-up sample. The p-value indicates if the difference is statistically significant.

Table B.3: Balance test comparing Control and Treated Villages

	Control	Treatment	p-value
Income	22517.415	25922.361	0.162
Income Source- Agri.	0.927	0.851	0.041
Income Source- Business	0.111	0.085	0.569
Income Source- Job	0.021	0.027	0.420
Income Source- Remit.	0.007	0.013	0.347
Income Source- Other	0.045	0.045	0.971
Age	39.533	39.128	0.760
Elementary Education	0.185	0.161	0.339
Higher Education	0.077	0.086	0.736
Informal Education	0.324	0.293	0.471
Univeristy Education	0.007	0.010	0.697
Secondary Education	0.268	0.261	0.820
Degree Centrality	4.619	5.082	0.025
Caste- Brahmin	0.052	0.088	0.565
Caste- Chhetri	0.216	0.296	0.581
Caste- Newar	0.460	0.395	0.755
Caste- Tamang	0.202	0.159	0.679
Caste- Dalit	0.059	0.020	0.289
Other	0.010	0.043	0.066

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

Notes: The balance tests compare characteristics of those in pure control and treated villages. The p-value indicates if the difference is statistically significant.

Table B.4: Correlation of Follow-up Survey Attrition with Treatment Status

	(1)
VARIABLES	Attrition
Spillover	-0.0449
	(0.0508)
Treated alone (T1)	-0.0501
	(0.0448)
Treated with a pair	-0.0540
	(0.0431)
Constant	0.425***
	(0.0335)
Observations	1,204
R-squared	0.002
Spillover==T1	0.925
T1=Paired	0.928
Spillover=Paired	0.837

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

Notes: This regression compares 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.

C Follow-up Results

Table C.1: Long-term effect on Business Outcomes

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Spillover	-0.0395**	-5,397	-2,034*	5,527
	(0.0158)	(18,721)	(1,006)	(13,573)
Treated alone (T1)	0.00591	8,815	1,093	21,594
	(0.0212)	(17,359)	(1,496)	(17,372)
Treated with a friend (T2)	-0.0213	20,074	-1,089	30,194
	(0.0185)	(21,340)	(1,156)	(20,586)
Treated with a friend + module (T3)	0.00457	34,543*	1,532	27,385
	(0.0221)	(19,425)	(1,594)	(17,783)
Constant	0.0395**	102,770***	2,034*	98,576***
	(0.0158)	(13,472)	(1,006)	(8,972)
Observations	750	734	750	589
R-squared	0.009	0.009	0.009	0.007
Spillover==T1	0.0127	0.205	0.0488	0.410
T1=T3	0.956	0.0358	0.817	0.752
T1=T2	0.115	0.503	0.162	0.731
T2=T3	0.137	0.513	0.0426	0.916

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. Agricultural profits, business investments, and agricultural investments are winsorised at 0.01% to exclude outliers. Standard errors are robust and clustered at the village level.

Table C.2: Long-term effect on Economic Outcomes

VARIABLES	(1) Monthly Income (Winsorised)	(2) Savings (Winsorised)	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Spillover	801.8	2,661*	0.0471	0.0502	0.0327
	(2,487)	(1,371)	(0.0426)	(0.0424)	(0.0621)
Treated alone (T1)	2,975	4,015***	0.107***	0.0874**	0.0633
	(2,563)	(1,281)	(0.0376)	(0.0381)	(0.0677)
Treated with a friend (T2)	3,317	2,783	-0.00151	0.0545	0.109*
	(3,034)	(1,800)	(0.0454)	(0.0338)	(0.0618)
Treated with a friend + module (T3)	5,785*	2,744*	0.0607	0.102**	0.149**
	(3,052)	(1,386)	(0.0443)	(0.0388)	(0.0650)
Constant	22,921***	3,401***	0.197***	0.0674***	0.446***
	(1,839)	(826.1)	(0.0277)	(0.0186)	(0.0490)
Observations	749	744	752	752	751
R-squared	0.007	0.015	0.010	0.012	0.011
Spillover==T1	0.353	0.341	0.151	0.417	0.607
T1=T3	0.264	0.373	0.374	0.730	0.111
T1=T2	0.888	0.526	0.0469	0.400	0.457
T2=T3	0.526	0.982	0.299	0.169	0.491

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

Notes: Monthly income and savings are winsorised at 0.01% to exclude outliers. Standard errors are robust and clustered at the village level.

Table C.3: Long-term effect on Economic Outcomes by Friendship Status

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Paired (Friend)	3,050	-865.1	-0.101*	-0.0172	-0.0150
	(2,313)	(1,774)	(0.0502)	(0.0359)	(0.0553)
Paired (Not Friend)	-1,434	-1,063	-0.0555	-0.0160	0.116*
	(2,242)	(1,362)	(0.0506)	(0.0437)	(0.0662)
Constant	25,425***	7,161***	0.281***	0.189***	0.457***
	(3,569)	(1,525)	(0.0538)	(0.0598)	(0.0877)
Observations	431	428	433	433	433
R-squared	0.006	0.001	0.009	0.002	0.015
Paired with friend==Nonfriend	0.135	0.874	0.416	0.967	0.0509

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

Notes: This regression treats individuals who were treated alone as the base category. Monthly income and savings are winsorised at 0.01% to exclude outliers. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

Table C.4: Long-term effect on Economic Outcomes by Centrality of Peer

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Paired (More Central)	-1,055	401.8	-0.120**	-0.0653	0.0705
	(3,541)	(1,509)	(0.0520)	(0.0508)	(0.0522)
Paired (Less Central)	1,661	-1,875	-0.0500	0.0121	0.0537
	(2,010)	(1,456)	(0.0439)	(0.0468)	(0.0624)
Constant	25,653***	6,288***	0.309***	0.214***	0.463***
	(4,123)	(1,407)	(0.0570)	(0.0673)	(0.0809)
Observations	435	432	437	437	437
R-squared	0.002	0.006	0.011	0.008	0.004
Paired with more central==Less central	0.512	0.0293	0.114	0.225	0.768

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

Notes: This regression treats individuals who were treated alone as the base category. Monthly income and savings are winsorised at 0.01% to exclude outliers. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

Table C.5: Long-term effect of the training by friendtype

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Friend x More Central	0.0339	4,185	895.3	-6,361
	(0.0428)	(15,507)	(3,055)	(33,366)
Friend x Less Central	-0.0413**	61,661**	-2,590*	10,520
	(0.0177)	(23,626)	(1,460)	(26,594)
Not Friend x More Central	-0.0433*	3,569	-4,275**	4,209
	(0.0229)	(17,616)	(1,731)	(28,105)
Not Friend x Less Central	-0.00285	10,343	1,175	10,131
	(0.0255)	(13,272)	(2,330)	(25,992)
Degree	-0.00560*	4,836	-676.8***	3,024
	(0.00289)	(3,115)	(232.3)	(4,205)
Constant	0.0773***	84,693***	6,939***	103,382***
	(0.0221)	(24,155)	(2,256)	(26,821)
Observations	432	424	432	318
R-squared	0.020	0.028	0.023	0.004
Friend x Less Central=Friend x More Central	0.0674	0.0281	0.159	0.645
Friend x Less Central=Non Friend x Less Central	0.0934	0.114	0.0899	0.992
Friend x Less Central=Non Friend x More Central	0.905	0.0441	0.0479	0.815
More Central==Less Central	0.443	0.0161	0.580	0.513
Friend=Non Friend	0.371	0.243	0.660	0.832

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

Notes: This regression treats Treatment 1 as the base category. Friend includes pairs that have social distance less than or equal to two and nonfriend includes pairs with social distance greater than two. Friend x More Central equals 1 when the individual the respondent is matched with is a friend and more central than the respondent. Agricultural profits, business investments, and agricultural investments are winsorised at 0.01% to exclude outliers. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

Table C.6: Long-term effect of the training by friendtype

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Friend x More Central	-1,467	827.4	-0.173**	-0.0182	-0.0938
	(3,502)	(2,384)	(0.0712)	(0.0653)	(0.0695)
Friend x Less Central	6,190*	-2,054	-0.0506	-0.0157	0.0390
	(3,552)	(1,724)	(0.0500)	(0.0404)	(0.0669)
Not Friend x More Central	-1,113	313.8	-0.0722	-0.0939	0.202**
	(5,159)	(1,670)	(0.0643)	(0.0549)	(0.0719)
Not Friend x Less Central	-1,856	-1,904	-0.0480	0.0343	0.0567
	(1,493)	(1,621)	(0.0536)	(0.0592)	(0.0727)
Degree	-52.19	158.8	-6.82e-05	-0.0108	0.0123
	(700.4)	(297.4)	(0.00815)	(0.00986)	(0.0119)
Constant	26,168***	6,321***	0.301***	0.218***	0.441***
	(4,478)	(1,440)	(0.0552)	(0.0693)	(0.0830)
Observations	431	428	433	433	433
R-squared	0.012	0.007	0.014	0.013	0.027
Friend x Less Central=Friend x More Central	0.173	0.153	0.0822	0.974	0.110
Friend x Less Central=Non Friend x Less Central	0.0291	0.921	0.964	0.317	0.794
Friend x Less Central=Non Friend x More Central	0.285	0.179	0.765	0.227	0.0637
More Central==Less Central	0.423	0.0153	0.0983	0.280	0.913
Friend=Non friend	0.244	0.897	0.378	0.662	0.0222

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

Notes: This regression treats Treatment 1 as the base category. Friend includes pairs that have social distance less than or equal to two and nonfriend includes pairs with social distance greater than two. Friend x More Central equals 1 when the individual the respondent is matched with is a friend and more central than the respondent. Monthly income and savings are winsorised at 0.01% to exclude outliers. We additionally control for the individual's own degree centrality. Standard errors are robust and clustered at the village level.

D Mechanisms

Table D.1: Impact of being trained with a Central Friend after controlling for Similarity

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Friend x More Central	0.0182	0.141	0.284	0.218**
	(0.189)	(0.101)	(0.169)	(0.0865)
Friend x Less Central	-0.0304	0.0758	0.0934	0.0546
	(0.119)	(0.0811)	(0.120)	(0.0995)
Not Friend x More Central	-0.177	0.0432	0.162	0.154*
	(0.132)	(0.0788)	(0.136)	(0.0861)
Not Friend x Not Central	-0.0974	0.0521	0.0925	0.111
	(0.186)	(0.0678)	(0.108)	(0.0926)
Similarity Index	0.274	-0.0484	-0.276	-0.153
	(0.211)	(0.116)	(0.236)	(0.132)
Constant	-0.129	0.741***	0.189	0.256***
	(0.158)	(0.0611)	(0.113)	(0.0867)
Observations	684	682	671	670
R-squared	0.007	0.010	0.013	0.007
Friend x Less Central=Friend x More Central	0.718	0.254	0.224	0.0272
Friend x Less Central=Non Friend x Less Central	0.638	0.530	0.601	0.144
Friend x Less Central=Non Friend x More Central	0.125	0.505	0.991	0.426
More central==Less Central	0.844	0.485	0.222	0.0557
Friend=Non Friend	0.174	0.122	0.429	0.928

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

Notes: This regression treats Treatment 1 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.

Table D.2: Effect of being matched with a Central Friend after controlling for Other Characteristics

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Friend x More Central	0.322	0.202**	0.437**	0.129
	(0.203)	(0.0905)	(0.195)	(0.105)
Friend x Less Central	0.292*	0.157*	0.301*	-0.0126
	(0.157)	(0.0865)	(0.166)	(0.121)
Not Friend x More Central	0.146	0.123	0.367*	0.0791
	(0.154)	(0.0922)	(0.200)	(0.118)
Not Friend x Not Central	0.198	0.126*	0.275	0.0348
	(0.131)	(0.0722)	(0.167)	(0.108)
Peer is Upper Caste	0.0469	0.0371	0.127*	0.0649
	(0.0873)	(0.0387)	(0.0740)	(0.0705)
Peer has higher Education	-0.0460	-0.0708	-0.0440	-0.0110
	(0.140)	(0.0445)	(0.0955)	(0.0682)
Peer has higher Income	-0.436***	0.0530	0.0104	-0.00201
	(0.0699)	(0.0398)	(0.0677)	(0.0523)
Peer is Married	0.0828	-0.0675	-0.266*	0.0255
	(0.120)	(0.0719)	(0.145)	(0.0959)
Peer is Younger	-0.0916	-0.110**	-0.274***	-0.108
	(0.108)	(0.0395)	(0.0814)	(0.0682)
Constant	-0.122	0.727***	0.164	0.247***
	(0.159)	(0.0632)	(0.114)	(0.0864)
Observations	684	682	671	670
R-squared	0.041	0.036	0.038	0.013
Friend x Less Central=Friend x More Central	0.806	0.374	0.336	0.0585
Friend x Less Central=Non Friend x Less Central	0.498	0.389	0.742	0.516
Friend x Less Central=Non Friend x Less Central	0.0568	0.522	0.601	0.202
More Central==Less Central	0.872	0.569	0.255	0.0864
Friend==Non Friend	0.142	0.176	0.548	0.979

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

Notes: This regression compares individuals within the paired treatment arms. Standard errors are robust and clustered at the village level.

Table D.3: Learning by whether pairs are friends or not

VARIABLES	(1) Knowledge Index	(2) Profit (Game)	(3) Profit (Business Plan)
Paired (Friend)	-0.0222	17.44	648,079
	(0.0857)	(94.25)	(741,374)
Paired (Non Friend)	-0.00239	-1.015	468,939
	(0.0624)	(89.86)	(590,952)
Constant	0.955***	299.0***	458,013**
	(0.0491)	(71.99)	(173,423)
Observations	688	688	687
R-squared	0.000	0.000	0.003
Friend==Non Friend	0.786	0.780	0.841

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

Notes: This regression treats Treatment 1 as the base category. Standard errors are robust and clustered at the village level.

Table D.4: Learning by type of pair

VARIABLES	(1) Knowledge Index	(2) Profit (Game)	(3) Profit (Business Plan)
Friend x More Central	-0.0208	0.996	867,927
	(0.109)	(96.62)	(880,202)
Friend x Less Central	-0.0402	19.49	413,558
	(0.0899)	(95.36)	(642,707)
Not Friend x More Central	0.0141	-3.338	489,008
	(0.0784)	(89.57)	(823,871)
Not Friend x Not Central	0.000488	8.157	509,319
	(0.0748)	(96.37)	(596,980)
Degree	0.0205**	11.41	80,176
	(0.00777)	(8.275)	(95,897)
Constant	0.846***	239.4***	37,455
	(0.0603)	(78.45)	(600,891)
Observations	684	684	683
R-squared	0.006	0.006	0.005
Friend x Less Central=Friend x More Central	0.852	0.568	0.219
Friend x Less Central=Non Friend x Less Central	0.527	0.769	0.943
Friend x Less Central=Non Friend x More Central	0.527	0.769	0.943
More Central==Less Central	0.802	0.825	0.391
Friend==Non Friend	0.607	0.871	0.627

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

Notes: This regression treats Treatment 1 as the base category. Standard errors are robust and clustered at the village level.

Table D.5: Correlation of Number of Contacts Pooled and Similarity with Peer

VARIABLES	(1) Number of links pooled in T3	(2) Number of links pooled in T3
Same age group	-0.155	
	(0.710)	
Same Caste	1.346**	
	(0.619)	
Peer has same education	0.187	
	(0.405)	
Same Income (Quartile)	-0.0959	
	(0.578)	
Same Marital Status	-0.474	
	(0.591)	
Quadratic Difference in Degrees		0.0188**
		(0.00891)
Constant	5.881***	6.264***
	(0.730)	(0.381)
Observations	216	206
R-squared	0.068	0.051

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

Notes: This regression compares individuals in Treatment Arm 3 with each other. Standard errors are robust and clustered at the village level.

Table D.6: Heterogeneous Treatment Effects by Number of Contacts Shared

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	-0.0975	0.0680	0.343***	0.288***
	(0.0971)	(0.0474)	(0.0891)	(0.0877)
Treatment 2	-0.0203	0.0940*	0.304***	0.355***
	(0.0822)	(0.0467)	(0.100)	(0.0907)
Treatment 3	0.0886	0.144**	0.418***	0.229*
	(0.184)	(0.0622)	(0.130)	(0.133)
T3 X Number of contacts pooled	-0.137	0.0251	0.00544	0.126
	(0.261)	(0.0862)	(0.155)	(0.183)
Constant	0	0.704***	0	0
	(0.0579)	(0.0283)	(0.0716)	(0.0659)
Observations	1,189	1,184	1,171	1,168
R-squared	0.002	0.019	0.038	0.037

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

Notes: Standard errors are robust and clustered at the village level.

Table D.7: Effect of Peer Risk Aversion

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	0.105	-0.00888	0.0982	0.106
	(0.176)	(0.0739)	(0.144)	(0.116)
Treatment 2 and 3	-0.0605	0.0541	0.0886	0.0904
	(0.0989)	(0.0505)	(0.0776)	(0.0949)
Risk Averse	-0.0230	-0.0560	-0.184*	-0.230**
	(0.0751)	(0.0577)	(0.104)	(0.104)
Treatment 1 X Risk Averse	-0.242	0.0387	0.136	0.136
	(0.183)	(0.0826)	(0.168)	(0.155)
Treatment 2 and 3 X Risk Averse	-0.0327	-0.0150	0.0210	0.0891
	(0.107)	(0.0638)	(0.123)	(0.126)
Constant	-0.0203	0.803***	0.226***	0.131*
	(0.0774)	(0.0480)	(0.0800)	(0.0736)
Observations	872	868	857	856
R-squared	0.013	0.026	0.073	0.063

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

Notes: Standard errors are robust and clustered at the village level.

Table D.8: Impact of friend type (by eigen vector centrality)

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Friend x More Central	0.146	0.0723	0.0261	0.0311
	(0.119)	(0.0557)	(0.0965)	(0.0733)
Friend x Less Central	0.107	0.0782	0.0492	0.0534
	(0.116)	(0.0459)	(0.0798)	(0.0785)
Not Friend x More Central	0.0894	0.00589	-0.0914	0.0958
	(0.119)	(0.0550)	(0.117)	(0.0762)
Not Friend x Not Central	-0.0358	0.0354	-0.00684	-0.0104
	(0.111)	(0.0367)	(0.0888)	(0.0751)
Eigen	3.827	1.605	0.706	0.863
	(6.193)	(1.188)	(2.159)	(1.562)
Constant	-0.149	0.756***	0.352***	0.270***
	(0.108)	(0.0501)	(0.0724)	(0.0546)
Observations	684	682	671	670
R-squared	0.007	0.012	0.004	0.004
Friend x Less Central=Friend x More Central	0.682	0.880	0.809	0.782
Friend x Less Central=Non Friend x Less Central	0.267	0.142	0.205	0.634
Friend x Less Central=Non Friend x More Central	0.895	0.236	0.451	0.284
More Central==Less Central	0.279	0.533	0.449	0.560
Friend==Non Friend	0.289	0.159	0.263	0.992

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

Notes: This regression treats Treatment 1 as the base category. Standard errors are robust and clustered at the village level.

E Robustness Checks

E.1 Short-Term Outcomes

Table E.1: Impact of the training (Main Effects)

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	-0.0682	0.0655	0.305***	0.256***
	(0.0922)	(0.0493)	(0.0854)	(0.0802)
Treatment 2	-0.0293	0.0724	0.212**	0.291***
	(0.0707)	(0.0456)	(0.0904)	(0.0837)
Treatment 3	-0.0150	0.120***	0.306***	0.236***
	(0.0719)	(0.0394)	(0.0695)	(0.0711)
Constant	-0.0611	1.050***	0.653***	0.830***
	(0.0526)	(0.0586)	(0.0952)	(0.126)
Observations	1,171	1,166	1,154	1,151
Number of groups	0	0	0	0
Treatment 1==2	0.671	0.902	0.359	0.644
Treatment 2==3	0.854	0.410	0.305	0.489
Treatment 1==3	0.516	0.238	0.987	0.756

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections.

Table E.2: Impact of the training (Paired vs Treatment 1)

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Treatment 1	-0.0678	0.0664	0.303***	0.256***
	(0.0936)	(0.0490)	(0.0847)	(0.0801)
Treatment 2 and 3	-0.0227	0.0996***	0.261***	0.265***
	(0.0653)	(0.0334)	(0.0684)	(0.0666)
Constant	-0.0605	1.040***	0.646***	0.831***
	(0.0545)	(0.0621)	(0.0963)	(0.125)
Observations	1,171	1,166	1,154	1,151
Number of groups	0	0	0	0
Treatment 1 == Pair	0.564	0.423	0.579	0.869

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections.

Table E.3: Impact of the training by friend type

VARIABLES	(1) Business Aspirations	(2) Ready to invest	(3) Business Index	(4) Take-up Index
Friend x More Central	0.170	0.0877	0.0775	0.113*
	(0.153)	(0.0583)	(0.130)	(0.0628)
Friend x Less Central	0.117	0.0444	-0.0723	-0.0317
	(0.102)	(0.0499)	(0.0731)	(0.0769)
Not Friend x More Central	-0.00549	0.00272	-0.0108	0.0509
	(0.0948)	(0.0600)	(0.132)	(0.0686)
Not Friend x Not Central	0.0568	0.0185	-0.0811	0.0300
	(0.116)	(0.0368)	(0.0825)	(0.0689)
Constant	-0.0990	1.109***	0.632***	0.714***
	(0.169)	(0.0506)	(0.119)	(0.108)
Observations	682	680	669	668
Number of groups	0	0	0	0
Friend X Less Central==Friend X More Central	0.691	0.366	0.267	0.0249
Friend X Less Central==Non Friend X Less Central	0.643	0.434	0.627	0.343
Friend X Less Central==Non Friend X More Central	0.170	0.450	0.877	0.238
More central==Less central	0.951	0.716	0.264	0.0811
Friend==Nonfriend	0.188	0.128	0.511	0.996

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections.

E.2 Long-Term Outcomes

Table E.4: Long Term Effects on Business and Agricultural Outcomes

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Treated alone (T1)	0.0124	9,678	1,861	25,977
	(0.0230)	(19,278)	(1,762)	(17,342)
Treated with a pair	-0.00399	27,499	689.8	32,332**
	(0.0194)	(19,670)	(1,234)	(14,426)
Constant	0.0651***	103,667***	3,830***	77,223***
	(0.0191)	(14,787)	(1,129)	(9,986)
Observations	726	709	726	571
Number of groups	0	0	0	0
Paired=Nonpaired	0.364	0.0584	0.450	0.726

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. The pure control group is used as the base category and we additionally control for an indicator for the spillover group. Covariates include income, source of income, age, education, caste, number of network connections. Profits and investments are winsorised at 0.01% to exclude outliers.

Table E.5: Long Term Effects on Economic Outcomes

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Treated alone (T1)	1,707	3,564***	0.103***	0.104***	0.0456
	(2,378)	(1,229)	(0.0378)	(0.0396)	(0.0673)
Treated with a pair	2,648	2,569*	0.0222	0.0864**	0.120**
	(2,319)	(1,456)	(0.0349)	(0.0338)	(0.0565)
Constant	18,323***	3,757***	0.184***	0.0836**	0.380***
	(2,515)	(1,216)	(0.0413)	(0.0339)	(0.0712)
Observations	725	720	727	727	726
Number of groups	0	0	0	0	0
Paired=Nonpaired	0.567	0.462	0.0525	0.630	0.132

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. The pure control group is used as the base category and we additionally control for an indicator for the spillover group. Covariates include income, source of income, age, education, caste, number of network connections. Monthly income and savings are winsorised at 0.01% to exclude outliers. Standard errors are clustered at the village level.

Table E.6: Long Term Effects on Business and Agricultural Outcomes by Friendship Status

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Paired (Friend)	-0.0107	39,441**	-1,214	3,798
	(0.0218)	(17,378)	(1,809)	(22,932)
Paired (Not Friend)	-0.0152	4,001	-581.2	5,685
	(0.0193)	(10,751)	(1,637)	(20,274)
Constant	0.0455***	111,586***	3,127**	120,169***
	(0.0167)	(12,832)	(1,487)	(15,136)
Observations	434	426	434	319
Number of groups	0	0	0	0
Paired with friend==Nonfriend	0.807	0.113	0.682	0.934

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections. Profits and investments are winsorised at 0.01% to exclude outliers.

Table E.7: Long Term Effects on Economic Outcomes by Friendship Status

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Paired (Friend)	3,035	-1,065	-0.103**	-0.0157	-0.0140
	(2,336)	(1,756)	(0.0485)	(0.0355)	(0.0559)
Paired (Not Friend)	-1,520	-1,272	-0.0608	-0.00938	0.109*
	(2,040)	(1,487)	(0.0487)	(0.0420)	(0.0613)
Constant	25,895***	7,416***	0.303***	0.155***	0.510***
	(1,792)	(977.8)	(0.0325)	(0.0330)	(0.0493)
Observations	433	430	435	435	435
Number of groups	0	0	0	0	0
Paired with friend==Nonfriend	0.0706	0.864	0.429	0.824	0.0446

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections. Monthly income and savings are winsorised at 0.01% to exclude outliers. Standard errors are clustered at the village level.

Table E.8: Long Term Effects on Business and Agricultural Outcomes by Centrality of Peer

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Paired (More Central)	-0.00379	-5,047	-1,127	-6,997
	(0.0217)	(11,118)	(1,705)	(22,486)
Paired (Less Central)	-0.0211	33,570***	-687.6	13,551
	(0.0195)	(10,291)	(1,739)	(18,121)
Constant	0.0455***	111,586***	3,127**	120,169***
	(0.0167)	(12,832)	(1,487)	(15,136)
Observations	438	430	438	323
Number of groups	0	0	0	0
Paired with more central==Less central	0.368	0.000639	0.776	0.238

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections. Profits and investments are winsorised at 0.01% to exclude outliers.

Table E.9: Long Term Effects on Economic Outcomes by Centrality of Peer

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Paired (More Central)	-1,141	-72.74	-0.120**	-0.0465	0.0570
	(3,011)	(1,747)	(0.0499)	(0.0430)	(0.0529)
Paired (Less Central)	1,656	-2,000	-0.0532	0.00980	0.0574
	(1,931)	(1,420)	(0.0413)	(0.0455)	(0.0619)
Constant	25,895***	7,416***	0.303***	0.155***	0.510***
	(1,792)	(977.8)	(0.0325)	(0.0330)	(0.0493)
Observations	437	434	439	439	439
Number of groups	0	0	0	0	0
Paired with more central==Less central	0.381	0.0729	0.0955	0.266	0.995

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. Covariates include income, source of income, age, education, caste, number of network connections. Monthly income and savings are winsorised at 0.01% to exclude outliers. Standard errors are clustered at the village level.

F Spillover Effects

In this section, we briefly discuss the effects of the intervention on individuals in the spillover group, one year after the intervention. As shown in Table F.1, we find that the spillover group is significantly less likely to open new businesses and invests significantly less in the new business as well. This is not only significantly different from those in the pure control group but also from those who were treated alone or in pairs.

Table F.1: Long-term effect on Business and Agriculture Outcomes.

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Spillover	-0.0395**	-5,397	-2,034*	5,527
	(0.0158)	(18,708)	(1,005)	(13,562)
Treated alone (T1)	0.00591	8,815	1,093	21,594
	(0.0212)	(17,347)	(1,495)	(17,357)
Treated with a pair	-0.00955	26,669	99.44	28,855**
	(0.0185)	(17,452)	(1,232)	(14,083)
Constant	0.0395**	102,770***	2,034*	98,576***
	(0.0158)	(13,463)	(1,005)	(8,965)
Observations	750	734	750	589
R-squared	0.007	0.008	0.005	0.007
Spillover=T1	0.0126	0.204	0.0487	0.409
Paired=Nonpaired	0.407	0.0773	0.530	0.680
Spillover=Paired	0.00923	0.0203	0.0212	0.0970

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. Agricultural profits, business investments, and agricultural investments are winsorised at 0.01% to exclude outliers. Standard errors are robust and clustered at the village level.

To understand this effect, we first ensure that this does not arise due to covariate imbalance. Tables F.2 presents the PDS-Lasso results where we additionally control for covariates including income, source of income, age, education, caste, number of network connections, and willingness to open businesses. We include baseline willingness to open businesses as a covariate (even though it is missing for 20% of the follow-up data) since we find that willingness to open businesses is 13 percentage points lower for the spillover group compared to the pure control group (at 10% significance). Even in this specification, we continue to find that the effect persists and the spillover group is less likely to open new businesses. However, at the same time, we find no effects on economic outcomes as seen in Table F.3.

We do not proceed with evaluating these findings any further. This is because only 3% of our follow-up sample have opened up new businesses and the investments in new business is also non-zero only for these individuals. This is a small set of individuals ($n = 23$) as a result of which drawing inferences using this variable does not seem sensible, especially when we do not any meaningful spillover effects on other important outcomes.

Table F.2: Long Term Effects on Business and Agricultural Outcomes

VARIABLES	(1) Opened New Business	(2) Agriculture Profits	(3) Investment in New Business	(4) Investment in Agriculture
Spillover	-0.0349**	-11,296	-1,694*	15,092
	(0.0145)	(22,128)	(898.2)	(15,979)
Treated alone (T1)	-0.00100	6,134	1,712	31,009
	(0.0256)	(22,381)	(2,194)	(20,816)
Treated with a pair	-0.0142	21,282	-41.52	28,226*
	(0.0175)	(21,768)	(1,220)	(14,486)
Constant	0.0362**	102,266***	1,771*	60,477***
	(0.0147)	(17,314)	(915.8)	(19,618)
Observations	593	579	593	471
Number of groups	0	0	0	0
Spillover=T1	0.0797	0.275	0.0772	0.489
Paired=Nonpaired	0.414	0.0996	0.307	0.892
Spillover=Paired	0.0350	0.0295	0.0314	0.404

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. The pure control group is used as the base category and we additionally control for an indicator for the spillover group. Covariates include income, source of income, age, education, caste, number of network connections, and willingness to open businesses. Profits and investments are winsorised at 0.01% to exclude outliers.

Table F.3: Long Term Effects on Economic Outcomes

VARIABLES	(1) Monthly Income	(2) Savings	(3) New Savings Account	(4) Taken a loan	(5) Commitment Savings
Spillover	-3,838	1,451	0.0550	0.0289	0.0549
	(2,496)	(1,392)	(0.0453)	(0.0435)	(0.0687)
Treated alone (T1)	1,913	2,866*	0.0802*	0.0937**	0.0744
	(2,676)	(1,473)	(0.0469)	(0.0417)	(0.0729)
Treated with a pair	2,562	1,049	0.0255	0.0775**	0.143**
	(2,488)	(1,285)	(0.0394)	(0.0387)	(0.0644)
Constant	24,047***	4,055***	0.187***	0.0663***	0.368***
	(1,805)	(955.6)	(0.0308)	(0.0198)	(0.0515)
Observations	592	588	594	594	593
Number of groups	0	0	0	0	0
Spillover=T1	0.0391	0.311	0.573	0.219	0.739
Paired=Nonpaired	0.734	0.233	0.314	0.684	0.143
Spillover=Paired	0.0255	0.681	0.523	0.298	0.147

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

Notes: The table reports treatment effects using post double selection Lasso (Belloni et al. 2014) to account for baseline covariates that may be corelated with treatment status. The pure control group is used as the base category and we additionally control for an indicator for the spillover group. Covariates include income, source of income, age, education, caste, number of network connections, and willingness to open businesses. Monthly income and savings are winsorised at 0.01% to exclude outliers. Standard errors are clustered at the village level.

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.

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.

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.

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.

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?’.

Cecchi-DiMeglio, P. (2023), ‘How leaders strategically enhance their networks for maximum impact’.
URL: *https://www.forbes.com/sites/paolacecchi-dimeglio/2023/11/29/how-leaders-strategically-enhance-their-networks-for-maximum-impact/*

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.

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.

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.

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.

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., 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.

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.