CHANGES IN SOCIAL NETWORK STRUCTURE IN RESPONSE TO EXPOSURE
TO FORMAL CREDIT MARKETS

Abhijit Banerjee
Emily Breza
Arun G. Chandrasekhar
Esther Duflo
Matthew O. Jackson
Cynthia Kinnan

Changes in Social Network Structure in Response to Exposure to Formal Credit Markets
Abhijit Banerjee, Emily Breza, Arun G. Chandrasekhar, Esther Duflo, Matthew O. Jackson,
and Cynthia Kinnan
NBER Working Paper No. 28365
January 2021
JEL No. D13,D85,L14,O12,Z13

ABSTRACT

Formal financial institutions can have far-reaching and long-lasting impacts on informal lending
and information networks. We first study 75 villages in Karnataka, 43 of which were exposed to
microfinance after we first collected detailed network data. Networks shrink more in exposed
villages. Links between households that were unlikely to ever borrow from microfinance are at
least as likely to disappear as links involving likely borrowers. We replicate these surprising
findings in the context of a randomized controlled trial in Hyderabad, where a microfinance
institution randomly selected neighborhoods to enter first. Four years after all neighborhoods
were treated, households in early-entry neighborhoods had credit access longer and had larger
loans. We again find fewer social relationships between households in early-entry neighborhoods,
even among those ex-ante unlikely to borrow. Because the results suggest global spillovers,
which are inconsistent with standard models of network formation, we develop a new dynamic
model of network formation that emphasizes chance meetings, where efforts to socialize generate
a global network-level externality. Finally, we analyze informal borrowing and the sensitivity of
consumption to income fluctuations. Households unlikely to take up microcredit suffer the
greatest loss of informal borrowing and risk sharing, underscoring the global nature of the
externality.

Abhijit Banerjee
Department of Economics, E52-540
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
banerjee@mit.edu

Emily Breza
Harvard University
Littauer Center, M28
1805 Cambridge Street
Cambridge, MA 02138
and NBER
ebreza@fas.harvard.edu

Arun G. Chandrasekhar
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
arungc@stanford.edu

Esther Duflo
Department of Economics, E52-544
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
eduflo@mit.edu

Matthew O. Jackson
Department of Economics
Stanford University
Stanford, CA 94305-6072
and CIFAR,
and also external faculty of the Santa Fe Institute
jacksonm@stanford.edu

Cynthia Kinnan
Department of Economics
Tufts University
8 Upper Campus Road
Medford, MA 02155
and NBER
cynthia.kinnan@tufts.edu

1. Introduction

Social networks are an important source of credit, insurance, information, advice, and other economic and non-economic benefits and often substitute for limited formal institutions.[1] But networks are not designed: they are the product of many decisions. In particular, as formal markets expand, the incentives to maintain or develop new relationships change. This could affect networks in unanticipated ways, potentially affecting even those who do not directly benefit from this expansion (Arrow, 2000; Putnam, 2000).

In this paper, we study how the introduction of formal lending institutions changes social networks, both empirically and theoretically. In our first empirical setting, we analyze how the introduction of microfinance (MF) affects network relationships in rural communities. We show that MF entry leads to a general shrinkage of the network, even among those whose characteristics make them very unlikely to be borrowing from the microfinance institution (MFI). In fact, it is they who, despite being *prima facie* uninvolved with microcredit, appear to be the most affected, losing a considerable number of relationships among themselves. Because existing models of network formation struggle to rationalize these patterns, we develop a new model that can explain these findings. Our model highlights spillovers stemming from the decision to socialize or not. We subsequently replicate these surprising findings in a second, independent empirical setting from a randomized controlled trial (RCT) offering microfinance to urban communities, demonstrating both the robustness of these findings and the fact that the loss in links persists even after microfinance is no longer available to these communities.

The challenge in ascertaining whether formal institutions change informal social structures is that it requires detailed data on networks of informal relationships, together with exogenous variation in access to formal institutions. Our two empirical contexts satisfy both requirements. First, we analyze the introduction of MF in rural Karnataka, India using detailed network panel data that we collected (Banerjee, Chandrasekhar, Duflo, and Jackson, 2013, 2019b) over six years in 75 villages. These villages were selected in 2006, prior to the first survey wave, when none of them had access to microfinance, but a microfinance institution, Bharatha Swamukti Samsthe (BSS) was planning to start operating in all of them. Between 2007 and 2010, BSS entered 43 of these 75 villages, which we call MF villages. However, a series of external crises halted BSS's expansion and the remaining 32 villages were not exposed to BSS prior to our Wave 2 survey, collected in 2012. We call these non-MF villages. We take advantage of this variation, along with our extremely detailed network data from the two waves (covering 16,476 households) to estimate the impact of MF on village network structure, using a difference-in-difference strategy.

---

[1]See, e.g., Udry (1994); Fafchamps and Lund (2003); Karlan, Mobius, Rosenblat, and Szeidl (2009); Beaman and Magruder (2012); Ambrus, Mobius, and Szeidl (2014); Blumenstock, Eagle, and Fafchamps (2016); Munshi and Rosenzweig (2016); Blumenstock and Tan (2016); Breza (2016).

Second, we replicate and extend the Karnataka findings, leveraging an RCT conducted in 104 neighborhoods in Hyderabad, India, using cross-sectional survey data that we collected (Banerjee, Duflo, Glennerster, and Kinnan, 2015a; Banerjee, Breza, Duflo, and Kinnan, 2019a). In the RCT, entry by an MFI (Spandana) was randomized to half of the study neighborhoods. Control areas began receiving access to Spandana two years later. But in 2010 Spandana suddenly ceased all operations, due to the same crisis that halted BSS expansion. We surveyed all households for a final time six years after initial entry when they had little or no access to microcredit. However, the households in the early entry neighborhoods had been exposed for twice as long when microcredit was shut down and had received much larger loans. We therefore estimate the impact of this differential access to microcredit using data we collected in this survey about each respondent's network relationships, as described below.

The advantage of the Karnataka setting is that we have (very) high-quality network data. We know details of link patterns between households as well as the nature of the link (e.g., financial, informational, social). Furthermore, it is a panel and so allows us to condition on pre-period network structure. However, the setting does not involve an RCT and therefore our identification relies on the difference-in-difference assumption being valid.

The Hyderabad dataset avoids this issue, since initial entry was randomized: Treatment neighborhoods had exogenously more cumulative access to microfinance than Control neighborhoods. It also serves as a validation because the hypotheses we test in this data come from the results of the Karnataka analysis, which were generated before we looked at the network data in Hyderabad. Finally, because the survey was fielded 6 years after initial entry and 4 years after the late-entry group received access to MF, this illustrates the extent to which these kinds of effects can be durable.[2] However, in Hyderabad we only have one cross-section of network information, and we only have partial network data. To supplement it, we collected "aggregated relational data" (ARD) and use the new methodology from Breza et al. (2020) to estimate features of the network. Our ARD survey asks each respondent to list their network relationships and to indicate how many of those individuals have a series of traits, (e.g., a household member who migrated abroad, a government job). Breza et al. (2020) and Breza, Chandrasekhar, McCormick, and Pan (2019) have shown that these responses contain sufficient information to identify the parameters of a network formation model which can then be used to estimate the key characteristics of the neighborhood network that we need for our analysis. Breza et al. (2020) and Breza, Chandrasekhar, McCormick, and Pan (2019) show that this method is an effective way of identifying effects on networks, comparable to the case where the researcher has full network data.

The impact of microfinance on network connections can potentially go in either direction. As a source of formal credit to poor, underbanked households, microfinance may reduce their dependence on social networks for informal credit and insurance. Moreover, the required weekly

---

[2]We do find durable, and even growing, impacts of early entry on pre-existing businesses.

repayment structure of microloans may reduce borrowers' liquidity and limit their capacity to lend small sums to their friends (Field, Pande, Papp, and Park, 2012). On the other hand, if households re-lend a part of their formal loans, microfinance could crowd in informal financial relationships.[3]

In both of our datasets we find that the introduction of microfinance crowds out social network relationships. The probability of a link between any two households declines by 11% ($p = 0.077$) in a MF village compared to a non-MF village in the Karnataka sample. This is robust to controlling for a rich array of baseline variables. We estimate an even larger effect in the Hyderabad RCT – a 22% decline ($p = 0.048$).

We then investigate how the changes in networks are distributed across two types of households: those who are likely to take up microfinance loans and those who are not. All of the channels described above suggest that microfinance might affect a borrower's willingness to maintain friendships, including with those who do not take up microfinance. However, prima facie (without any sort of externality or spillover), one would not expect effects on pairs (or groups) of households that are both unlikely to take up microfinance. If anything, one would have expected links between these households to be *strengthened* in microfinance villages, since they might be losing access to the households that get microfinance but still have needs to borrow and lend.

To look at this question empirically, we need to be able to compare those who are more or less likely to take up microfinance in MF villages/neighborhoods to those in a non-MF village who would have been comparably likely to take up microfinance had it been available in their village/neighborhood. To this end, we use a random forest model to classify households in all villages into two groups based on whether they would have a high (H) or low (L) likelihood of joining microfinance if it were offered in their village.

We begin with the Karnataka panel by looking at the difference between MF and non-MF villages in the probability that two Ls who were linked in Wave 1 continue to be linked in Wave 2. Because L households have a low propensity to borrow from microfinance, they are unlikely to experience any direct impact. The surprising result is that LL links decline as much as LH links and more than HH links in MF villages relative to non-MF. An LL link that exists in Wave 1 in a MF village is 5.8pp ($p = 0.003$) less likely to exist in Wave 2 compared to a similar link in a non-MF village; this decline is, if anything, greater than the decline in HH links (the p value for the difference in coefficients is 0.086 without controls and 0.28 with). Similarly, LL links are less likely to form in MF villages, again, even less so than HH links.

The cross-sectional data from the Hyderabad RCT delivers consistent results. Treated MF neighborhoods have 0.6pp (22%) fewer LL links than control neighborhoods ($p = 0.023$), and there is no evidence of a greater decline in LH or HH links.

---

[3]Kinnan and Townsend (2012); Field, Pande, Papp, and Park (2012); Feigenberg, Field, and Pande (2013); Vera-Cossio (2019) find evidence consistent with re-lending bank and credit cooperative loans in Thailand.

We then examine the evolution of links that form triangles. In the Karnataka sample, we find that it is the *LLL* triangles that are most likely to disappear in MF villages compared to non-MF villages. In MF villages, *LLL* triangles are 7.8pp ($p = 0.008$) more likely to have at least one link broken than in non-MF villages, more than any other type. The difference is greatest and most significant between *LLL* and *HHH*, but even *LHH* are less likely to break than *LLL* (by 5.4pp, $p = 0.072$). *LLL* triangles are also more likely to entirely disappear in MF villages, and the difference from all of the other types of triangles is significant. In the Hyderabad data, we also find that we are significantly less likely to observe a *LLL* triangle in treatment than control villages.

Thus, we observe loss of links even among people least likely to be involved in microfinance and in parts of the network that do not directly involve a connection with *H*s. These findings suggest that models of purely local externalities are unlikely to be able to explain our results.

Strikingly, even though the direct impact of microfinance is likely to be on financial links, the same patterns also emerge when we analyze information (i.e., advice-giving and -receiving) links. This suggests that there is contagion from one type of relationship to others.

These types of spillovers, both across types of links and across types of households, are *prima facie* inconsistent with models of network formation where the decision to form a link only depends on the payoff to the two parties forming the link and these payoffs only depend on the characteristics of the parties involved in the link and no one else. We briefly sketch a set of these models that are standard in the literature in Online Appendix D: these include mutual consent models of directed search, stickiness in dropping or forming links, and local payoff externalities.[4]

A potential explanation (also described in more detail in Online Appendix D) for why pre-existing *LL* links also drop in large numbers comes from a model of network formation with local payoff externalities. Many models of network formation focus on payoff externalities (see, e.g., Jackson et al. (2012); Mele (2017).Specifically, an *LL* link may be partly sustained by its shared connections to an *H* that was directly impacted by microfinance. However, the result that *LLL* triangles are at least as likely to be affected as triangles involving *H*s rules this out as a sole explanation.

We develop a new model of network formation that can explain why links between the *L*s might break at least as much as other links. The model we build comes from a simple idea. In the model, old relationships are maintained and new ones are formed when people socialize in an "undirected" way. A stylized interpretation is that people show up at the town square, or a local tea shop, to "hang out" and socialize. Seeing their current friends keeps those relationships intact, and meeting new people sometimes results in new relationships. People who do not show

---

[4]See, e.g., Jackson and Wolinsky (1996); Dutta and Mutuswami (1997); Bala and Goyal (2000); Currarini and Morelli (2000); Jackson and Van den Nouweland (2005); Herings et al. (2009); Boucher (2015); Watts (2001); Jackson and Watts (2002); Christakis et al. (2010a); König et al. (2014); Currarini et al. (2009, 2010); Cabrales et al. (2011); Canen et al. (2017).

up at the town square lose old relationships and form fewer new ones. We describe this as a model of undirected search.

This gives rise to a distinct *network-level* externality, because the returns to socializing depend on who else is socializing. Holding fixed the valuation of a certain link or groups of links, the fact that, in equilibrium, others are not searching can have global effects on network density and topological structure. For example, L types value HL links, and thus care about how H types socialize. Therefore, if microfinance changes the socialization of H types, that changes incentives for L types to socialize, which in turn affects the incidence of LL links. Specifically, access to microcredit might reduce both the demand and supply of informal loans by H types, but the Hs becoming less willing to lend can have a larger negative impact on Ls than on Hs, which leads to less socializing by Ls. As Ls socialize less there is a larger relative drop in LL links. A simple extension of the model to account for the formation of triads (triangles) generates similar results for LLL relationships.

This model matches the patterns we observe in the data, in particular the spillovers to the relationships between L types. It also predicts that there should be spillovers across different types of relationships, since it is the same town square where people also form other relationships such as advice relationships.

Given that we see network connections shrinking, a natural question is whether we see changes in downstream outcomes, such as borrowing or the volatility of consumption. Indeed, consistent with the disappearance of the LL links, we find in both settings that the L households, after the introduction of microfinance, borrow relatively less from informal sources in MF compared to non-MF villages.

Finally, in the Hyderabad sample we can directly measure the impact of increased microfinance exposure on consumption smoothing for high- and low-borrowing propensity households. This is possible because we have detailed panel information on both income and consumption at the household level. We find that, in areas exposed to microfinance, households with high propensity to use microfinance see little change in their consumption smoothing compared to those in areas not exposed to microfinance. However, households with low propensity to use microfinance see a large and significant worsening of their consumption smoothing compared to those in areas not exposed to microfinance, which is consistent with the network and informal borrowing impacts.

Our research on how exposure to formal financial institutions affects social and economic networks is related to some important recent and ongoing work. Feigenberg et al. (2013) find that participation in microcredit creates tighter social relationships among group members. Binzel, Field, and Pande (2013) and Comola and Prina (2014) explore whether and in what ways financial interventions also end up affecting those households' networks.[5]

---

[5]Specifically, Binzel et al. (2013) look at network effects in a randomized roll-out of branches of a new financial intermediary in India. Their focus is on whether individuals are less likely to make transfers to their friends in a non-anonymous dictator game after being exposed to the financial institution. Comola and Prina (2014) study

In recent work, Heß, Jaimovich, and Schündeln (2020) also examine how policy interventions affect network structure but in the context of a community-driven development initiative (CDD). The initiative provided a very large disbursement—one half of annual per capita income— *per household* in each treatment village, and villagers had to collectively decide which projects to execute. Heß et al. (2020) collected a cross-section of network data in 2014 and, like us, document declines in network density and closure, which in their case are generated by political maneuvering and elite capture. A key difference between CDD and microcredit is that the injection of the former is massive and at the community level whereas micro-loans are both smaller in size and are only suitable for a small subset of the community; so the general equilibrium effects on network structure come from very different sources and for different reasons.[6]

Our study contributes to and extends this line of inquiry. A main lesson from our paper is the presence of significant and widespread spillovers in network formation across types of people and types of relationships, which is indicative of a global network externality. We also use this evidence to build and argue for a new model of network formation that highlights the fact that social networks are not designed but result from the decentralized decisions of individuals. As our empirical results highlight, in such an environment, a shift in the incentives of one group of people to form links can have substantial (negative) effects on other parts of the network and groups that they ignore when choosing their own behavior.[7]

The remainder of the paper is organized as follows. In Section 2, we describe the setting, network data collection, the classification of households into H and L types using a random forest algorithm, and sample statistics. Section 3 presents our empirical results. Motivated by the data, in Section 4 we develop a new dynamic model of network formation that is consistent with it and discuss why four standard models from the literature are inconsistent with the data. Section 5 presents impacts on informal borrowing and the capacity for households to smooth consumption. Section 6 concludes.

2. SETTING, DATA AND SAMPLE STATISTICS

2.1. Setting.

2.1.1. *Karnataka (India).* In 2006, the microfinance organization, BSS, provided us with a list of 75 villages in Karnataka in which they were planning to start lending operations. The villages were spread across 5 districts of the state of Karnataka in India. Prior to BSS's entry, these villages had minimal exposure to microfinance.

---

how individuals' social networks change when randomly assigned to receive a savings account in Nepal. Their focus is on post-intervention expenditure spillovers, taking into account network change due to the exposure to the savings account.

[6]Nonetheless, our model could still be useful in understanding the effects in such interventions.

[7]See Jackson (2003) for background on inefficiencies in network formation. Here we see general, network-level externalities. Of course, this does not mean that microcredit should be discouraged, but only any welfare analysis needs to take into account the potential for spillovers.

Six months prior to BSS's entry into any village, in 2006, we conducted a baseline survey in all 75 villages. This survey consisted of a village questionnaire, a full census that collected data on all households in the villages, and a detailed follow-up survey fielded to a subsample of adults.

By the end of 2010, BSS had entered 43 villages that were not randomly assigned by us, but rather selected by the bank. We have anecdotal reasons to believe that the choice was not systematic: BSS planned to enter all of the villages but slowed down and ultimately stopped expanding during the Andhra Pradesh (AP) microcredit crisis (see Breza and Kinnan (2018) for background on that crisis).

2.1.2. *Hyderabad (India).* In 2006 Spandana –a large microfinance institution– randomly chose 52 of 104 neighborhoods in Hyderabad (at the time the capital of Andhra Pradesh, a State neighborhing Karnataka, in South India) to enter. After two years, the remaining 52 neighborhoods received access in mid-2008. The short- and medium-run impacts of randomized access to microfinance in this setting are studied in Banerjee et al. (2015a). The AP microcredit crisis also impacted Spandana and its lending activities in Hyderabad. In 2010, all of the households in the Hyderabad sample faced simultaneous withdrawal of microcredit in response to an ordinance halting microcredit loans (this also means they did not need to repay existing debt). A third round of data collection was done in 2012, with a sample of 5744 households. At the time of the 2012 data collection, the treatment neighborhoods had been exposed to microcredit for 6 years (4 years active lending) and the control neighborhoods had been exposed for 3.5 years (with 1.5 years active lending). Network data was collected during this third round.

The early treatment neighborhoods had greater microfinance access overall. Because microfinance borrowers typically receive larger loans each time they borrow, microcredit supply is increasing in the length of exposure. Banerjee et al. (2019a) show that two years after the control group received access, households in treated neighborhoods still had 14% more contemporaneous microfinance borrowing and 43% more cumulative microfinance borrowing over the preceding three years (Banerjee et al., 2019a). However, since nobody had access to microfinance at the time of our network survey, any changes to network structure that we pick up must be the result of the extra exposure to microcredit before it was shut down some two years before our survey. In other words, the effect persists despite there being no differences in contemporaneous participation in microcredit.

2.2. **Data.**

2.2.1. *Karnataka.* To collect the network data,[8] we asked adults to name those with whom they interact in the course of daily activities. In Wave 1, collected in 2006, we have the full

---

[8]The Wave 1 data are described in detail in Banerjee, Chandrasekhar, Duflo, and Jackson (2013) and publicly available at `http://economics.mit.edu/faculty/eduflo/social`. The Wave 2 data will be available upon publication.

village census (enumerating every individual in every household in every village and some basic household characteristics) and network data from 46% of households per village. In Wave 2, collected in 2012, in addition to taking the full village census again, we have network data from 89.14% of the 16,476 households. This means that we have network data in Wave 1 on 70.8% of the links and in Wave 2 on 98.8% of the links when we build the undirected, unweighted graph that we study.[9] For the network analysis, we concentrate on households that are present in both waves and only look at objects (e.g., potential links or potential triads) where we are able to discern in both waves whether the structure exists or does not exist.

We have data about 12 different types of interactions for a given survey respondent: (1) whose houses he or she visits, (2) who visits his or her house, (3) relatives they socialize with, (4) non-relatives they socialize with, (5) who gives him or her medical help, (6) from whom he or she borrows money, (7) to whom he or she lends money, (8) from whom he or she borrows material goods (e.g., kerosene, rice), (9) to whom he or she lends material goods, (10) from whom he or she gets important advice, (11) to whom he or she gives advice, (12) with whom he or she goes to pray (e.g., at a temple, church or mosque).

Using these data, we first look at the financial network (a union of (6-9) above) as well as the informational network ((10-11) from above). After demonstrating that links across both categories change in similar ways, we aggregate the network data as follows. We construct one network for each village, at the household level, where a link exists between households if any member of either household is linked to any other member of the other household in at least one of the 12 ways. We assume that individuals can communicate if they interact in any of the 12 ways, so this is the network of potential communications. The resulting objects are undirected, unweighted networks at the household level.

We also asked, in both Wave 1 and Wave 2, for households to give us a list of all outstanding loans that they have taken, the sources of these loans (e.g., family member, friend, microfinance institution, self-help group, money lender) and their terms. We use this to create a panel to study changes in borrowing patterns.

In our analysis we look at all households who existed in Wave 1 (and in Wave 2 as well). This involves those who remained and those who split. We match households who split in Wave 2 to their Wave 1 counterpart. 11% migrated out, though this is not differential by microfinance exposure, and 4.8% Wave 2 households in-migrated (which we cannot use in the panel) or split off from existing households (as children reach adulthood), again not differential by microfinance exposure.[10]

---

[9]The 70.8% figure is calculated as follows. Because we consider a non-directed graph, we learn about the existence of a link when either participating node is sampled. Therefore for arbitrary nodes A and B, Pr(sample either A or B) $= 1 - (1 - 0.46)^2 = 0.708$.

[10]Note that when we construct the panel, our sample of potential links ij conditions on the event that either i or j was surveyed in period 1 in the case of link tables (and the analogous construction for triangles). Thus, we can be sure that we are studying the evolution of links or triangles in a way that is not plagued by sampling issues (Chandrasekhar and Lewis, 2014).

2.2.2. *Hyderabad.* The Hyderabad analysis draws on three waves of data. These data are also utilized in Banerjee et al. (2015a) and Banerjee et al. (2019a). The first round of data collection was conducted in late 2007 - early 2008, 15-18 months after microfinance was made available in the treatment group. Following this first wave, the control group also received access to microfinance in May 2008. A second round of data collection was conducted in mid-2010 to examine longer-term impacts of access to microfinance; coincidentally, this wave took place just before the AP Crisis, mentioned above. Finally, in 2012, approximately two years after the AP Crisis, a third wave of data collection took place. All three waves collected information about household composition, income, consumption/expenditure, borrowing (from microfinance and from other sources) and entrepreneurship. For the third wave only, we also measured aspects of households' social networks. However we could only collect partial network data across the 104 neighborhoods in Hyderabad, so instead we collected Aggregated Relational Data (ARD). Because we collected this information only in the 2012 Wave 3 data, the majority of our analysis uses Wave 3 only; an exception is the analysis of consumption smoothing which leverages the panel nature of the data.

Specifically, an average of 55 households in every neighborhood in the Hyderabad sample were surveyed and asked a set of network questions. First, respondents were asked how many links they had within the neighborhood (eliciting their degree) along three dimensions: financial, social and informational.[11] This is the directly solicited part of the network information. Second, respondents were asked 9 ARD questions of the form "How many individuals from your neighborhood do you know who have trait X?" For instance, traits include "How many other households do you know where there are 5 or more children?" and "How many other households do you know where any member is a permanent government employee." Supplemental Appendix E.1 details both types of survey questions. Third, we asked each sampled household whether they possessed each of the ARD traits.

We use the method of Breza et al. (2020) to leverage ARD data to estimate key network characteristics. We give a more detailed description of the algorithm in Supplemental Appendix E.2. ARD counts the number of links an agent has to members of different subgroups in the population. The basic idea is that by combining this information with a model of network formation, one can estimate which possible networks would have generated this sample. Specifically, we assume a "latent distance" model, where the probability of a connection depends on individual heterogeneity and the (inverse) distance between pairs of nodes in a latent social space (Hoff, Raftery, and Handcock, 2002). Basically, by triangulating the information about how many nodes of each type a given respondent knows, we can learn a lot about where the individual is likely to be located in the latent space. This allows us to estimate a distribution over possible network configurations that could have generated this sample information. We

---

[11]Specifically, we asked who individuals would go to and who would come to them for borrowing basic goods (cooking gas, a small amount of cash, etc.), advice (e.g., on health or education), and socializing (watching TV).

can then generate graphs from this distribution and compute network statistics for each generated graph. For many applications, that information about the potential networks is enough to draw relevant conclusions. [12] Here we are mainly interested in the frequency of different types of relationships, which is easily recovered from ARD. Note that the way we elicited the ARD means that we only have information about one single type of link encompassing all dimensions of interaction, both financial and non-financial.[13]

2.3. **Sample Statistics and Covariate Balance.** Table 1, Panel A shows Wave 1 (baseline) network characteristics by treatment status in Karnataka, while Panel B shows Wave 1 household demographics. The networks are sparse: the average density is 11.9%. The average clustering coefficient (the percent of cases where two of a household's friends are themselves friends) is 0.33. Finally, these networks have short distances: the average closeness (the mean of the inverse of path lengths, with 0 taken for nodes on different components) is 0.379.[14] In general, the network structure and demographic variables look quite similar between the microfinance villages and non-microfinance villages at baseline. However, the MF villages are larger, on average, than the non-MF villages.[15] We further examine baseline balance on an expanded range of demographic and village-level covariates in Online Appendix Table C.1, Panel A.

Next we turn to Hyderabad. Recall that since we have an endline cross-section, we only measure the network characteristics after the intervention to test for balance. In Online Appendix Table C.2, we show the means of select network characteristics in control neighborhoods. In this urban sample, the networks are even more sparse than Karnataka; the average degree is 6.0, for an average neighborhood size of approximately 200 households. Average clustering and closeness are also smaller. In Table 1 Panel C, we present sample statistics and tests of covariate balance using a set of predetermined neighborhood and household characteristics. Given that the introduction of microfinance was randomized in the Hyderabad sample, the covariates are

---

[12]For example, Breza et al. (2020) shows that ARD data can replicate the results of Breza and Chandrasekhar (2019) as well as if the entire network was observed. We also validate ARD in the Hyderabad dataset. Specifically, in the surveys, we directly measured support – the likelihood that for any link, there exists a third person who has a relationship with both nodes. We validate ARD by comparing the estimated measure of support using the ARD algorithm with the directly elicited survey measure. In this way, we can show that the ARD estimate leads to the identical conclusions.

[13]The ARD algorithm requires information on the total population of each neighborhood. Unfortunately, we were unable to collect this information for 15 of the 104 neighborhoods. Therefore, in specifications that use population as an input (e.g., density, graph-level clustering, link-level analyses), we drop these neighborhoods from the analysis.

[14]In order to deal with the fact that we sampled data in Wave 1, we compute average density among the sampled households in Wave 1, comparing the share of realized links relative to potential links when we fully observe the potential link. We compute the clustering coefficient among the subgraph induced by restricting to sampled households in Wave 1, since that is centered around the true parameter. It is also worth noting that the correlation among the different link types (specifically multiplexing of information and financial links) is 0.638.

[15]In Online Appendix M, we show that our main results are robust to allowing for differential trends by functions of village size.

balanced in treatment and control. Again, Online Appendix Table C.1 includes an expanded set of pre-determined covariates.

2.4. **Classifying Nodes as H and L.** In order to study heterogeneity in effects by propensity to participate in microfinance, we need to identify which households would have had taken out microfinance loans in the non-microfinance villages or neighborhoods, had BSS or Spandana entered those villages. To do this, we use a random forest model to classify an individual's propensity to take up microfinance as a function of baseline characteristics, in the presence of microfinance. We can then use this classification exercise to predict which individuals in the entire sample (treatment and control) have a high propensity to borrow.

We begin with the Karnataka setting. One obvious determinant of microfinance take-up is from the BSS rules: only households with a female in the age range 18-57 were eligible for microfinance. Also, certain households were identified by BSS as a "leader" household and were specifically informed about the product.[16] Therefore leaders, or people close to them in the network, are more likely to have heard of the microfinance opportunity and have taken it up (Banerjee et al., 2013). We estimate the random forest model based on household demographics and network characteristics from the microfinance villages on a training sample of 7199 households and then validate the method on a testing sample of 2399 households. The features are as follows: (1) a dummy for the household being a BSS leader, which are households with an individual that the microfinance institution would approach when entering a village; (2) a dummy for whether the household has a female of eligible age (below 57), which was a requirement to be able to participate in microfinance; (3) the average closeness (mean of inverse of network distance) to leaders, which is relevant, because as in Banerjee et al. (2013), those who are closer to leaders should be more likely to hear of microfinance; (4) the average closeness (mean of inverse distance) to same-caste leaders, because interactions within caste are more likely and therefore should influence the likelihood of being informed; and (5) the share of same-caste leaders in the village. The details of the estimation algorithm, implemented choices, and quality are presented in Appendix B.

Turning to the Hyderabad setting, the strategy is similar, though Spandana did not have such clear rules for selecting borrowers. Thus, we consider 19 predictors of a household's take-up of Spandana, including demographic characteristics of the household (such as characteristics of the household head and his spouse, the number of women and children in the house, whether the household owns a business) as well as demographic data for the village (such as literacy rate, village population, total number of businesses in the village). We again use random forests, training a model on 2520 households and then validating the model on a testing sample of 1080 households.

---

[16]The BSS definition of leader was defined by occupation (e.g., teachers, self-help group leaders, shopkeepers), so we can identify them similarly in MF and non-MF villages.

We then apply the classifier to both microfinance villages (or neighborhoods) and non-microfinance villages (or neighborhoods) to classify each household as H or L (high or low likelihood of joining microfinance).

A major advantage of using random forests is that they naturally allow for non-linearities and potentially complex interactions between characteristics that could drive microfinance take-up. If the likelihood of being a microcredit taker or not is very non-linear in the characteristics, then random forest provides a very sensible and flexible approach. Alternatives such as logistic regression would not be able to handle such interactions and non-linearities without typically introducing a very high dimensionality of interaction terms. A related advantage of random forest comes from its value in identification. Because random forests allow for classification via a complicated non-linear function of the network and relation to leadership positions, in the Karnataka data where we have baseline network data, we can control smoothly for network position and network position interacted with post. Therefore, unobservables correlated smoothly with network parameters are unlikely to drive the Karnataka results.

Random forest classification does have a few downsides. First, because of the highly non-linear structures that can arise, the actual mapping from characteristics to classification are less interpretable than with logistic regression models. This is not a major problem in our case, since we are more interested in prediction than interpretation. Second, if the true underlying data-generating process has log-odds that are linear in parameters, then the random forest may overfit. Therefore, for robustness, we also present our main results in Appendix K using logistic regression to classify households into H and L types for both Karnataka and Hyderabad. In Appendix Section B.5, we show that random forest outperforms logit in both samples in terms of typical classification quality metrics. Further, all of the results are replicated in the Karnataka sample, where logit classification is at least comparable to the random forest. In Hyderabad, logit classifies types much worse.[17]

Table 2 presents some summary statistics from the classification exercise. In Panels A and B we look at Karnataka data. There are notable differences between H and L households. Although none of these features were used in the estimation, we find that H households are much more likely to be SC/ST, have smaller houses in terms of room count, are much less likely to have a latrine in the household, and are much less likely to have an RCC (reinforced concrete cement) roof, all of which suggests that they tend to be poorer. Finally, we see that H households have somewhat larger degrees than L households, and the composition exhibits homophily: H types have a lower number of links to L types and a higher number of links to H types. Finally H households are more eigenvector central in the network, which is not

---

[17]This can be seen even on first principles: the amount of borrowing by a high type versus a low type is no different in microfinance neighborhoods under the logit classifier ($Microfinance \times H + H \approx 0$, Online Appendix Table K.11, Column 1).

surprising given that they were selected to be close to leaders, who themselves tend to be more central. In Section 5.1 we show that indeed H types borrow considerably more than L types in microfinance villages. H types borrow Rs. 1,787 ($p < 0.001$) more than L types, indicating that the classification performs well.

Panels C and D turn to the Hyderabad data, and look only at the non-microfinance villages. In Panel C, in contrast to Karnataka, we find no significant difference between H and L households in their demographic characteristics. Turning to network characteristics, in Panel D we see, like in Karnataka, that H types have fewer links to L types, more links to H types, and are more central. Again, in Section 5.1 we show that one year after microfinance entered the treated neighborhoods, H types had considerably more microcredit than L types in early microfinance neighborhoods (Rs. 8,773, $p < 0.001$).

3. Changes in Networks

How does exposure to microfinance change networks? We begin with a discussion of how the overall structure of social networks are affected and then the effects on different types of bilateral links as well as triads.

3.1. **Effect on the total number of links.** We first look at how introducing microfinance affects the overall structure of the village social networks. In the Karnataka data, where we have a panel but no randomization, we use a difference-in-differences framework:

$$y(\mathbf{g}_{vt}) = \alpha + \beta \text{Microfinance}_v \times \text{Post}_t + \gamma \text{Microfinance}_v + \eta \text{Post}_t + \delta' X_v + \epsilon_{vt},$$

where $y(\cdot)$ computes the density of the network $\mathbf{g}_{vt}$ for village v in period t, the average closeness (the mean of the inverse distance between all pairs), or clustering. The density is the percentage of links a random household has to all other households in the village, so it measures how well-connected the village is on average.[18] The distance in the network is the (minimum) number of steps through the network it takes to get from one household to another. In models where favors, transactions, or information travel through the network, higher distance or lower closeness (the inverse) means that the movement of such phenomena through the network is slower. Finally, clustering is the share of a household's connections that are themselves connected. Economic models of network formation identify clustering as an important feature to sustain cooperation. In the Hyderabad data, where we only have endline data, we run a cross-sectional specification. X_v is a vector of control variables, which varies according to the specification.

Table 3, panel A presents the results for Karnataka. Columns 1-3 present the result for network density, columns 4-6 for clustering, and 7-9 for closeness. The first column in each category (columns 1, 4, 7) present a simple difference in differences specification. The second column in each specification (2, 5, 8) adds to that a vector of baseline controls interacted

---

[18]Note that density is directly related to average degree—it is proportional to average degree scaled by $n - 1$.

with $textPost_t$ as well as the controls interacted with treatment and Post$_t$. These controls include share of upper-caste households, number of households in the village, network density, share of households in self-help groups, share Hindu, share with a latrine in the house, share that own the household, share that have electricity and share that are leaders. We add these because differences in the size of the village, caste composition, or the wealth distribution could potentially have differently-evolved networks even without introduction of microfinance. While the entry of BSS does not seem to correlate with much of anything beyond village size, we include these controls to ensure that they do not drive the results. Finally, the third column in each specification (3, 6, 9) includes village fixed effects as well as controls for the baseline value of the outcome variable interacted with Post, to allow for differential time trends by baseline network feature. Because we only have 150 observations but many controls (up to 18 controls and their interactions before adding the fixed effects), we use the double post-LASSO (DPL) procedure (Belloni and Chernozhukov, 2009; Belloni, Chernozhukov, and Hansen, 2014a,b) to select the controls.[19]

We find that exposure to microfinance leads to a drop in density by about 1.2-1.3pp relative to a mean of 11.4% in non-microfinance villages in Wave 1 (columns 1-3, $p = 0.077$ in column 3 for example). This is an 11% drop in density. We don't find any detectable effect of microfinance on the clustering of the village. This is true irrespective of whether controls are used. Without controls we find a significant decrease in the average closeness (column 7, $p = 0.02$), corresponding to a 0.53 standard deviation effect. However, this loses significance in columns 8 and 9 with the inclusion of controls ($p = 0.19$, $p = 0.21$, respectively).

Panel B turns to the Hyderabad data, which uses an endline cross-sectional dataset rather than a panel, but takes advantage of the random selection of neighborhoods to treatment. There, we run the following specification.

$$y(\mathbf{g}_{vt}) = \alpha + \beta \text{Microfinance}_v + \delta' X_v + \epsilon_{vt},$$

Our vector of controls X_t are demographic characteristics of the household and the village, the same ones used for classification of H and L. We again use DPL to select the control variables. We find that there is a 22% decline in density ($p = 0.086$ without DPL and $p = 0.048$ with DPL). We do not find meaningfully significant results on clustering or closeness.

Thus, in both settings, we find a reduction in the density of the network.

3.2. **How are links affected by microfinance?** In this subsection, we explore how microfinance exposure affects the formation of links across types of households – our Hs and Ls.

Bilateral links can be of three types: HH, LH, and LL. Let $g_{ij,v,t}$ be an indicator for whether a link is present between households i and j in village v in wave t. Letting LH_{ij} be an

---

[19]Because the double post-LASSO procedure does not select all of the village fixed effects, this means that we can include the fixed effects and an indicator for microfinance in the same regression.

indicator for pair consisting of one low type and one high type, and analogously for HH_{ij} etc., the regressions we run take the form

$$g_{ij,v,2} = \alpha + \beta MF_v + \beta_{LH} MF_v \times LH_{ij,v} + \beta_{HH} MF_v \times HH_{ij,v}$$
$$+ \gamma_{LH} LH_{ij,v} + \gamma_{HH} HH_{ij,v} + \delta' X_{ij,v} + \epsilon_{ij,v,2},$$

where $X_{ij,v}$ includes a vector of flexible controls (a polynomial) for centrality of both nodes, demographic variables (caste and a number of wealth proxies including number of rooms, number of beds, electrification, latrine presence, roofing material), all variables that are used in the random forest classification, and then interactions of all of these variables with the microfinance dummies (the control variables finally included are chosen by DPL).

The idea behind identification is that the classification type, H or L, is a complex, non-linear function with many interaction terms of a subset of the features described above. As such, we can still smoothly control for them and allow the control to vary by whether the village is exposed to microfinance or not. This allows us to control for the potentially differential effect of microfinance exposure on households that are demographically distinct and located differently in the network under the maintained assumption that these effects can be capture by linear uninteracted terms. The coefficients of interest capture whether being in a microfinance village differentially affects the evolution of a link among types classified as HH, HL, and LL, conditional on all the characteristics above and their interaction with MF. We also present regressions without any controls whatsoever to demonstrate that the results are robust to the presence or absence of these detailed controls. Altonji, Elder, and Taber (2005) show that if the results do not change when we introduce more and more controls, this provides some support for the view that unobservables are not spuriously driving the results.

We run these regressions in two samples: the set of ij such that $g_{ij,v,1} = 1$ (in this case we ask whether pre-existing links break) and the set where of ij such that $g_{ij,v,1} = 0$ (so the link doesn't exist in the first period), in which case we ask about the probability of a new link forming in Wave 2.

Table 4 presents the link-level results for any type of relationship in the Karnataka data. In columns 1-2 we focus on the set of links existing in Wave 1 and in columns 3-4 we focus on the set of unlinked nodes in Wave 1. Columns 1 and 3 include no controls whatsoever. Columns 2 and 4 introduce the set of controls variables and their interaction with MF, selected by double-post LASSO. The key coefficients for testing the hypotheses are the coefficients on Microfinance, which captures the effect on the omitted category, LL, links, as well as Microfinance$\times LH$ coefficients and Microfinance$\times HH$, which ask whether the effects are different for these types of links, compared to LL. Columns 1 shows that LL links break significantly more in MF villages relative to non-MF villages. Specifically, they are 5.8pp less likely to exist in Wave 2 ($p = 0.002$), on a base of 48.2% in non-MF villages. The decrease in LH links is very similar,

while the HH links are somewhat less likely to disappear (and although the difference with the LL links is not statistically significant, it is worth noting that on average the HH links are not statistically more likely to break in MF villages than in non-MF villages). The results are robust to the inclusion of control variables.

Columns 3 and 4 present similar results for link formation. LL links are 2.3pp less likely to form in microfinance villages on a base of 7.5% in non-microfinance villages ($p = 0.006$). Again, the effect is comparable for LH links, while it is less pronounced for HH links. All of these results are robust to smoothly controlling for the centrality of nodes involved as well as demographic controls and their interactions with microfinance.

The relative changes in network structure in the microfinance villages shed light on network formation. The fact that the LH links break might reflect the fact that the Hs are no longer interested in maintaining their links with the Ls now that they have an alternative source of credit, but the fact that LL links are equally likely to break is more surprising, especially since the Ls should have a stronger incentive to hold on to their mutual links precisely because they no longer have access to the links with the Hs. [20]

We turn to the Hyderabad data in Table 5. In this case, we have only cross-sectional information on networks so we cannot condition on pre-period link status. Therefore we run the regression in the sample of any possible link ij. The "microfinance" coefficient identifies the effect on LL links (the omitted category) and captures a combination of link formation and link destruction. Column 1 includes only the randomization strata as controls, while column 2 additionally allows for any of the household or village level controls used in the random forest classification to be included. In column 1, we find a 0.6 percentage point (on a base of 2.7 percent) decline in the probability that an LL link exists in microfinance neighbhorhoods relative to non-microfinance neighborhoods ($p = 0.023$). We cannot reject that the estimates for LH and HH are the same, but they are imprecisely estimated. The estimates are quite similar in column 2, after adding controls.

We next unpack these findings for financial links (those that we anticipate would directly be affected by the credit injection) versus information links. Table 6, Panel A presents the results in the Karnataka data, where columns 1 and 2 consider the evolution of financial links, while columns 3 and 4 consider non-financial links. Columns 1 and 3 restrict to links of each type that existed in the Wave 1 data, while columns 2 and 4 restrict to pairs of individuals that were not linked in Wave 1. The patterns are strikingly similar across financial and information links, which is evidence of multiplexing. In fact, for information links, we find that the disappearance of HH links is significantly smaller than that of LL links.

---

[20]In Online Appendix M, we show that these impacts are robust to differential trends by village size interacted with link type. We also show a specification interacting treatment with each of the controls.

In the Hyderabad data, recall that the link-level information analyzed in Table 5 is constructed using ARD, which defines a link as a relationship occuring on any dimension (information, advice, or financial), so we cannot run an exactly parallel specification. However, we did collect supplemental, node-level information on relationship type that allows us to explore whether microfinance affects financial and non-financial links differentially. Panel B of Table 6 presents household-level regressions where the outcome variable is the number of financial or non-financial links, and the regressors are MF and $MF * H$ (with and without control variables). The main effect of MF identifies the effect of microfinance exposure for L households. It is is negative and highly significant on both the number of financial links and the number of non-financial links. The interaction effect, Microfinance$\times H$, identifies the differential effect of microfinance access for H households, those with high propensity to borrow. The interaction effects are positive and significant for all outcomes considered. The total effect for H households is in fact positive for all outcomes.

As in Karnataka, non-financial links have similar patterns as financial links, consistent with multiplexing. And while H households appear to hold steady or even gain links in response to microfinance exposure, the L households clearly and unambiguously lose links.

3.3. **Group Relationships.** In the link-level analysis we show that LL links — relationships between two individuals who experience minimal, if any, direct impacts from microfinance — are at least as likely to be affected as relationships involving H types. One natural place to look first to try to understand this result is local payoff externalities: does the decline in LL links stem from these housesholds' links to other H households who join microcredit?

Bloch, Genicot, and Ray (2008), Ambrus, Mobius, and Szeidl (2014), Jackson, Rodriguez-Barraquer, and Tan (2012) all propose models where contract enforcement requires groups of nodes rather than simple pairs. In Jackson, Rodriguez-Barraquer, and Tan (2012), for example, two households seeking to exchange favors may not have enough bilateral interaction to be able to sustain cooperation in isolation. However, if they both have relationships with some other households in common, then the relationships can all "support" each other and provide incentives to cooperate: if someone fails to cooperate with one of their friends, then beyond losing that relationship, they can also lose relationships with all of the other friends that they had in common.

Our network data exhibits such groupings, with the rates of a group of nodes being collectively linked far exceeding the rate to be expected if decisions were made independently (see Online Appendix G.2). These interdepencies in link formation may explain the impact of microfinance on LL links. If there are payoff externalities, L types might value an LL link more when there is a third node involved. The introduction of microcredit could destabilize these structures. In groups that are composed of both L and H types, it could be the case that microfinance directly causes LH links to break, which in turn spills over to adjacent LL links in the same group. In this world, groups only composed of L links should experience minimal impacts.

The direct testable implication is that, if we focus on triangles that existed in Wave 1, we should see a larger decline in triangles involving at least one H than in LLL triangles.

We use the following regression specification to test this hypothesis:

$$y_{ijk,v,2} = \alpha + \beta MF_v + \beta_{LHH} MF_v \times LHH_{ijk,v} + \beta_{LLH} MF_v \times LLH_{ijk,v} + \beta_{HHH} MF_v \times HHH_{ijk,v}$$

$$+ \gamma_{LHH} LHH_{ijk,v} + \gamma_{LLH} LLH_{ijk,v} + \gamma_{HHH} HHH_{ijk,v} + \delta' X_{ijk,v} + \epsilon_{ijk,v,2},$$

where $y_{ijk,v,2}$ is either a dummy for whether the triangle ijk exists in Wave 2 ($g_{ij,v,2} g_{jk,v,2} g_{ik,v,2} = 1$) in some specifications or whether any link in the former triangle exists in Wave 2 ($g_{ij,v,2} + g_{jk,v,2} + g_{ik,v,2} > 0$) in other specifications. The vector $X_{ijk,v}$ includes flexible controls for centralities of households, demographic characteristics previously described for all households, all classification variables used in the random forest model and the interactions of all of these variables with microfinance. As before, we present regressions with and without control variables.

Table 7 presents the results in the Karnataka data. In column 1, we find that all triad types (except the HHH) break faster in microfinance villages relative to non-microfinance villages. However, LLL triads that existed at baseline are the most likely to break in microfinance villages. Specifically, they are 7.8pp more likely to dissolve in microfinance relative to non-microfinance villages ($p = 0.008$). LHH triangles and HHH triangles are both statistically less likely to dissolve than LLL triangles ($p = 0.07$ and 0.028 respectively). The results are similar, if less precise, with control variables. Similarly, in column 3, we see that, out of formerly-linked triangles, we are more likely to see that none of the links survive for LLL triangles in MF villages (-8.5 pp, p value ≤ 0.001), and that this is significantly less likely to occur for LLH, LHH, and HHH triangles.

Table 8 presents the Hyderabad results and measures whether microfinance affects the event that a given set of three households are all linked (recall that we do not have baseline data, so we cannot condition on pre-existence). Because the likelihood of any potential triangle being fully linked is low (approximately 0.01%), we scale all regressors by 1,000, for readability.[21] Although the results are noisier than in Karnataka, we find once again that LLL triangles are negatively affected by microfinance: in column 1, we are 60% less likely to see any LLL triangle in MF neighborhood ($p = 0.089$). The effect is the same for LHH and smaller for LLH (p=0.12). The only difference is that HHH are the triangles who appear to be most likely to be missing in MF villages.

In summary, while there is evidence for interdependencies in the persistence of network relationships in our data, we find that microfinance affects LLL triangles particularly strongly. This suggests that models of local externalities are unlikely to be able to explain our results. The next section proposes an alternative model that can better rationalize them.

---

[21]This sparsity of groups of triangles also implies that a pooled cross-sectional analysis will largely reflect new link creation rather than existing link maintenance.

4. A Model

In this section we present a new model of network formation that is consistent with what we see in the data. The model provides new perspectives, both on how opportunities to interact arise and on how multiplexing works.

The externalities in the model do not take direct forms like value of friends-of-friends, or support, or other link interactions that are often used to incorporate externalities; they instead arise more globally in the search process by which people make and maintain friendships (see the discussion in Appendix D). We present the model for links and then describe how it can be extended to cover triads. As the model is useful beyond the current setting of microfinance, we describe it in a general form and then specialize to the two-type (H, L) microfinance case.

4.1. **Types and Utilities.** There are n individuals, indexed by $i, j... \in \{1, \ldots, n\}$. Agent i has a type θ_i from a type set Θ. Let $v_{\theta\theta'}$ denote the base benefit that an agent of type θ gets from a relationship with an agent of type θ'. For example, in our context, this can come from borrowing and lending activities, as we discuss in more detail below.

The realized utility from a relationship also involves an idiosyncratic noise term ε_{ij} that i gets from being friends with j. This could be personality compatibility, or some other benefits. Thus, an agent i gets a value $v_{\theta_i\theta_j} + \varepsilon_{ij}$ from a connection with j, where ε_{ij} is distributed according to an atomless distribution F.

A useful expression is

$$\mathrm{E}^+[v] = \mathrm{E}\left[v + \varepsilon_{ij} \,|\, \varepsilon_{ij} > -v\right] = v + \frac{\int_{-v}^{\infty} \varepsilon_{ij} dF}{\int_{-v}^{\infty} dF},$$

which denotes the expectation of $v + \varepsilon_{ij}$ conditional the value of $v + \varepsilon_{ij}$ being positive. This is the expected utility that an agent gets from a relationship with base value v, conditional upon being willing to form the friendship.

An agent of type θ then has an expected utility from $d_{\theta\theta'}$ friends of type θ' of

(4.1) $$\sum_{\theta' \in \Theta} d_{\theta\theta'} \mathrm{E}^+[v_{\theta\theta'}].$$

4.2. **Efforts and Link Formation.** Each agent chooses an effort $e_i \in [0, 1]$, which represents the amount of time they spend socializing to form and maintain links. In the case of the villagers, this could be time spent in the town square or tea shop, where they meet with other villagers.[22] As will become evident, our model is meant to capture both link formation and link maintenance.

Two agents i and j who have chosen efforts e_i and e_j have probability proportional to $e_i e_j$ of meeting. The model therefore rules out "directed search" since the probability of meeting

---

[22]This is a useful and conventional modeling device. See Currarini, Jackson, and Pin (2009, 2010); Cabrales, Calvó-Armengol, and Zenou (2011); Canen, Jackson, and Trebbi (2017) for other models where socialization takes effort and there is random meeting.

is independent of the agent's type, conditional on their effort. Time goes in periods $t \in \{0, 1, 2, \ldots\}$.

Let $g^t \in \{0,1\}^{n \times n}$ be the adjacency matrix representing network at time t. If $g_{ij}^{t-1} = 1$, then those agents were connected in the last period and they keep their friendship if they meet with each other during time t — keeping the relationship requires seeing each other. Therefore agents i and j keep their friendship with probability $e_i e_j$ and lose it with probability $1 - e_i e_j$. If $g_{ij}^{t-1} = 0$, then agents i, j were not connected in the last period and form a friendship with probability

$$e_i e_j \left(1 - F(-v_{\theta_i \theta_j})\right) \left(1 - F(-v_{\theta_j \theta_i})\right).$$

This is the probability that they meet *and* they both find the friendship of positive value – a friendship requires mutual consent in our model.

Thus, the efforts of agents do two things: they maintain old relationships by continuing an interaction but also allow them to meet new people.

4.3. **Steady-State Equilibrium.** A *steady-state equilibrium* is a list of efforts $(e_\theta)_{\theta \in \Theta}$, and a corresponding set of expected degree levels $(d_{\theta\theta'})_{\theta\theta' \in \Theta^2}$ such that e_θ maximizes each agent's expected utility, and the expected degree levels are in steady state as generated by the efforts.[23] We prove in the Appendix A that in all equilibria all agents of the same type choose the same action, and that the equilibrium is unique, provided that costs of effort are large enough.

The requirement that degrees are in steady state and generated by the efforts can be formally represented as follows. Let $n_{\theta\theta'}$ denote the number of agents of type θ' with whom an agent of type θ could potentially form friendships. If $\theta' \neq \theta$ then this will generally be the number of agents of type θ',[24] while if it is of type θ then it will less by one to account for the agent herself.

Out of those agents only an expected fraction of $\left(1 - F(-v_{\theta\theta'})\right)\left(1 - F(-v_{\theta'\theta})\right)$ will ever be friends with an agent of type θ, given the mutual consent requirement. Thus, let

$$m_{\theta\theta'} = n_{\theta\theta'} \left(1 - F(-v_{\theta\theta'})\right) \left(1 - F(-v_{\theta'\theta})\right).$$

This is the effective size of the pool of agents of type θ' with which an agent of type θ will be friends over time.

Degree at the end or beginning of a period is then the maintained relationships plus the new ones formed:

$$d_{\theta\theta'} = e_\theta e_{\theta'} d_{\theta\theta'} + (m_{\theta\theta'} - d_{\theta\theta'}) e_{\theta'} e_\theta,$$

---

[23]We solve the model in terms of steady-state and expected values, but it will be clear from the analysis that one can also do this in terms of realized values. The equilibrium will still be unique for sufficient costs of effort, the complementarities still apply in the same manner, and the equilibria have the same comparative statics. The complication is that strategies then need to be specified as a function of more than just type, as the realized noise terms then matter. Since those noise terms add no insight, we work with the more transparent version of equilibrium.

[24]It could also incorporate some other taboos or restrictions, for instance if some types simply are not permitted to form relationships, which would be captured by the vs.

which simplifies to

$$d_{\theta\theta'} = m_{\theta\theta'} e_\theta e_{\theta'} = n_{\theta\theta'} \left(1 - F(-v_{\theta\theta'})\right)\left(1 - F(-v_{\theta'\theta})\right) e_\theta e_{\theta'}.$$

Thus, in steady state, the degree is then proportional to the number of available agents of the other type, weighted by the probability that there is a mutual compatibility, and by the socializing efforts.

The expected utility of an agent involves the benefits from relationships, the costs of socialization, $-\frac{1}{2}c_\theta e_\theta^2$, as well as a benefit just from from socializing, $u_\theta e_\theta$. An agent may get some value from going to the town square or getting tea, etc., independently of who else is there.

Overall this leads to a utility of

$$V_\theta\left(e_\theta\right) = \underbrace{u_\theta e_\theta - \frac{1}{2}c_\theta e_\theta^2}_{\text{base socializing benefit and cost of effort}} + \underbrace{\sum_{\theta'\in\Theta} \mathrm{E}^+[v_{\theta\theta'}] d_{\theta\theta'} e_{\theta'} e_\theta}_{\text{expected maintenance of existing friendships by effort}}$$

$$+ \underbrace{\sum_{\theta'\in\Theta} \mathrm{E}^+[v_{\theta\theta'}] \left(m_{\theta\theta'} - d_{\theta\theta'}\right) e_{\theta'} e_\theta}_{\text{expected new friendships from effort}}$$

Using the expressions for $m_{\theta\theta'}$ and $d_{\theta\theta'}$, we can write this as

$$V_\theta\left(e_\theta\right) = u_\theta e_\theta - \frac{1}{2}c_\theta e_\theta^2 + \sum_{\theta'\in\Theta} \mathrm{E}^+[v_{\theta\theta'}] n_{\theta\theta'} \left(1 - F(-v_{\theta\theta'})\right)\left(1 - F(-v_{\theta'\theta})\right) e_{\theta'} e_\theta.$$

If we take $u_\theta > 0, c_\theta > 0$ for all θ and $\mathrm{E}^+[v_{\theta\theta'}] \geq 0$ for all θ, θ', then an equilibrium requires that:[25]

$$e_\theta = \min\left\{1, \frac{1}{c_\theta}\left(u_\theta + \sum_{\theta'\in\Theta} \mathrm{E}^+[v_{\theta\theta'}] n_{\theta\theta'} \left(1 - F(-v_{\theta\theta'})\right)\left(1 - F(-v_{\theta'\theta})\right) e_{\theta'}\right)\right\}.$$

4.4. **Equilibrium Existence and Some Comparative Statics.** This is a game of strategic complements, and for such games equilibria exist and form a complete lattice.[26] If $u_\theta = 0$ for all θ, then there exists a corner equilibrium in which all agents exert 0 effort. To examine the more interesting case, we presume that $u_\theta > 0$ for all agents, so that agents gain some utility from socializing regardless of the connections they form from it. In this case, for high enough costs of socialization there exists a unique equilibrium which has the property that there are spillovers from a change in the preferences of any type on the effort choices of all other types. The following result is proven in Appendix A.

---

[25]These come from the first order conditions, capped by the bound on efforts. Second order conditions are $-c_\theta$ and so are negative. Thus, these conditions are also sufficient.

[26]For instance, see Van Zandt and Vives (2007).

PROPOSITION **1.** *Let $u_\theta > 0, c_\theta > 0$ for all θ. For sufficiently large $c_\theta > 0$'s, there is a unique equilibrium and it is stable[27] and interior ($0 < e_\theta < 1$ for all θ), and agents of the same type take the same efforts. In addition, if $\mathrm{E}^+[v_{\theta\theta'}] > 0, n_{\theta\theta'} > 0$ for each θ, θ',[28] and $v_{\theta\theta'}$ is reduced for some $\theta\theta'$ (holding all other parameters constant), then $e_{\theta''}$ decreases for all θ'', and $d_{\theta''\theta'''}$ decreases for all $\theta''\theta'''$.*

The characterization of equilibrium is as follows. Let u be the $|\Theta|$-dimensional vector with entries $\frac{1}{c_\theta}u_\theta$ and E be the $|\Theta| \times |\Theta|$ matrix with θ, θ' entries

$$\frac{1}{c_\theta}\mathrm{E}^+[v_{\theta\theta'}]n_{\theta\theta'}\left(1 - F(-v_{\theta\theta'})\right)\left(1 - F(-v_{\theta'\theta})\right).$$

Then the unique equilibrium is given by

$$e = (I - E)^{-1}u,$$

which we show is well-defined for large enough costs in the Appendix A.

A major implication of the proposition is that a reduction in the returns from any type of relationship decreases *all* efforts and degrees. The decrease in value $v_{\theta\theta'}$ for some $\theta\theta'$ directly affects their efforts, decreasing those. Then, given the strict strategic complementarities, there is then a decrease in other efforts; and the feedback can lead to a substantial drop in all efforts.

Note that the relative drops in efforts and degrees predicted in Proposition 1 are not necessarily ordered across groups: degree can fall most among groups of nodes that experience no direct decline in link valuation (e.g., *LL* links in the case of microfinance). The intuition is that if marginal benefits to *L*s from connecting to *H*s are particularly high, then when *H*s drop effort, payoffs for *L*s from effort can drop even more than for *H*s, leading to an even larger effect on *L*s. See Online Appendix N for a simulation demonstrating this phenomenon.

4.5. **Externalities in Network Formation.** Even though our model does not include direct externalities in payoffs between links, the network-formation process still exhibits significant external effects since agents' decisions to form links (their effort levels) affect others' potential payoffs and their network formation decisions (e.g., agent 3 putting in less effort lowers the efforts of both agents 1 and 2 and thus the chance that 1 and 2 are linked, even though the potential $1 - 2$ link has nothing to do with agent 3).

This makes a point beyond the current setting: network formation can be inefficient not simply because of direct externalities in relationships, which is how it is usually modeled,[29] but also because meeting people requires effort, and this naturally generates strategic complementarities and substantial externalities.

---

[27]Slight changes in efforts lead to best reply dynamics that converge back to the equilibrium.

[28]All that is needed for this result is that this holds for a cycle of θ and θ' that include all types. Note also that $\mathrm{E}^+[v_{\theta\theta'}] > 0$ does not require that all people form links, just that there is a non-zero probability that any two types could find a high enough noise term to form a friendship.

[29]This class of models can incorporate, *inter alia* risk sharing, information sharing, and network support. For references see Jackson (2003, 2008).

4.6. **Specializing to Microfinance.** We next present a specialized case of the model to an-
alyze how microfinance changes incentives for socialization. We first present the steady-state
conditions for the case of two types, H and L. We then discuss how the introduction of micro-
finance maps to the model parameters, specifically the $v_{\theta,\theta'}$s.

4.6.1. *Two Types.* We now specialize the model to the case of two types: $\Theta = \{H, L\}$. Let
λ be the share of H types in the population. In this case, a steady-state is a solution to the
equations:

$$c_H e_H^\star = u_H + \mathrm{E}^+[v_{HH}] (\lambda n - 1) (1 - F(-v_{HH}))^2 e_H^\star$$
$$+ \mathrm{E}^+[v_{HL}] (1 - \lambda) n (1 - F(-v_{HL})) (1 - F(-v_{LH})) e_L^\star,$$

$$c_L e_L^\star = u_L + \mathrm{E}^+[v_{LL}] ((1 - \lambda) n - 1) (1 - F(-v_{LL}))^2 e_L^\star$$
$$+ \mathrm{E}^+[v_{LH}] \lambda n (1 - F(-v_{LH})) (1 - F(-v_{HL})) e_H^\star,$$

$$d_{HL} = ((1 - \lambda) n) e_H^\star e_L^\star (1 - F(-v_{HL})) (1 - F(-v_{LH})),$$

$$d_{LH} = d_{HL} \frac{\lambda}{1 - \lambda},$$

$$d_{HH} = (\lambda n - 1) (e_H^\star)^2 (1 - F(-v_{HH}))^2,$$

$$d_{LL} = ((1 - \lambda) n - 1) (e_L^\star)^2 (1 - F(-v_{LL}))^2.$$

The equilibrium vector of efforts ($e_\theta^\star$s) and network structure (vector of $d_{\theta\theta'}$s) are determined
by this system.

4.6.2. *Application to Microfinance: An Example.* How does the entry of microfinance affect
these parameters? Here, we present a rationalization for the payoffs based on informal bor-
rowing and lending. In particular, let the values $v_{\theta\theta'}$ be interpreted as "financial" payoffs from
borrowing and lending. This tells us how $v_{HH}, v_{HL}, v_{LH}, v_{LL}$ change in response to Hs getting
microcredit.

Lending produces a net return of r, which represents the effective expected interest rate from
informal loans less the opportunity cost of funds. Borrowing leads to a return net of repayment
of b, which represents the difference between the interest rate charged by a network "friend"
and the opportunity cost of foregoing that loan (e.g., losing the money or borrowing at some
higher rate from a money lender, etc.). Generally, we expect $b > 0$ and $b > r$,[30] as otherwise
such relationships make little sense. Whether r is positive or negative is not obvious since there

---

[30]The limited evidence we have on peer-to-peer lending suggests that markups on loans to friends tend to be
small, potentially even negative. b by contrast ought to be substantial and positive.

NETWORK CHANGE   24

are clearly social expectations to help out friends in need (which could make r negative), and may depend on context.

A household can be in one of three states of the world: they have money to lend, they need to borrow, or neither. An H household has a probability α_H of having money to lend and a probability $\beta_H \leq 1 - \alpha_H$ of needing to borrow, and with the remaining probability $1 - \alpha_H - \beta_H$ neither occurs. There are similar probabilities α_L and β_L for the L types.

The base payoff to an agent of type $\theta \in \{H, L\}$ of being matched to agent of type $\theta' \in \{H, L\}$ is then

$$v_{\theta\theta'} = \alpha_\theta \beta_{\theta'} r + \beta_\theta \alpha_{\theta'} b.$$

As in the general model, we assume that expected utility is additive across all relationships (Equation 4.1) and that pairwise payoffs $v_{\theta\theta'}$ are independent of other friendships.[31]

The introduction of microfinance changes these parameters. There are several likely channels by which the introduction of microfinance affects the payoff parameters of H types from linking to others. Access to microcredit might impact both the demand and supply of informal loans by H types. If access to microcredit substitutes for informal loans, then we would expect β_H to decrease. If alternately, the weekly required repayments are burdensome to households, they may have to cut back on lending smaller sums to others in the village and may even start borrowing small amounts to repay the loans when cash is short, leading to a decline in α_H and perhaps an increase in β_H. In addition, if there are complementarities between formal and informal loans because receiving a MF loan allows the household to overcome a non-convexity[32], β_H could go up. In contrast, if re-lending of formal credit to network partners is common, a type H may have a probability $\alpha'_H \geq \alpha_H$ of being able to lend once she gets access to microfinance. Her probability of needing to borrow may also go down to $\beta'_H \leq \beta_H$, if microfinance loans are a substitutes for network credit. In any case, we maintain that the Ls' needs for borrowing and lending are unaltered by the introduction of microfinance. Let

$$\Delta\beta_H = \beta'_H - \beta_H \quad \text{and} \quad \Delta\alpha_H = \alpha'_H - \alpha_H$$

be the changes in the probabilities that the H types have borrowing and lending needs after microfinance. By our previous assumption, $\Delta v_{LL} = 0$.

Let $\Delta_{\theta\theta'}$ denote the resulting change in $v_{\theta\theta'}$. To get a feeling for how this depends on $\Delta\alpha_H$ and $\Delta\beta_H$, note that for small values of $\Delta\alpha_H$ and $\Delta\beta_H$, we get approximations

$$\Delta_{HH} = (\alpha_H \Delta\beta_H + \beta_H \Delta\alpha_H)(r + b) \qquad \Delta_{LL} = 0$$
$$\Delta_{HL} = \alpha_L \Delta\beta_H b + \beta_L \Delta\alpha_H r \qquad \Delta_{LH} = \alpha_L \Delta\beta_H r + \beta_L \Delta\alpha_H b.$$

As we describe above, the arrival of microfinance may impact the valuations through a range of mechanisms. Different mechansims imply that $\Delta\beta_H$ and $\Delta\alpha_H$ could each be positive or

---

[31]We make this assumption to highlight our key mechanism of interest – that there can be spillovers to LL links that might be even bigger than spillovers to LH links despite not loading it into the payoff.

[32]See Banerjee et al. (2019a).

negative, making it very hard to say anything general about the signs of these expressions. For example, consider the special case in which $\alpha_L = \alpha_H$, $\beta_L = \beta_H$ and $\alpha_H \Delta \beta_H + \beta_H \Delta \alpha_H = 0$. In this case $\Delta_{HH} = 0$. Meanwhile, as discussed above, we expect $b - r$ to be positive. Then Δ_{HL} should be positive whereas Δ_{LH} should be negative as long as $\Delta \beta_H > 0$ and $\Delta \alpha_H < 0$.[33] Given that different valuations move in different directions, the prediction of how the different types should respond is unsigned.

However one obvious special case is when both α_H and β_H go down. In this case, as long as both b and r are positive, it is easy to see that all of v_{HL}, v_{LH} and v_{HH} must go down. Thus, Proposition 1 applies, implying that $e_H^\star, e_L^\star, d_{HH}, d_{HL}, d_{LH}, d_{LL}$ should all fall. However, the relative declines in degree across groups are unsigned. In Online Appendix N, we use simulations to show that it is indeed possible that LL links may be most affected, given the decreased effort of Hs as well as the Hs' lower propensity to want to link with Ls given their decreased borrowing needs. If this payoff drop is large enough for Ls, given their original benefits from connections to Hs, this can lead to a larger drop in Ls' efforts, and hence in LL links.

4.7. **Extensions of the model.** The simple model can be extended in several directions.

First, the model is solved in steady-state. Adding a population of unlinked (say "new-born") agents to the population of the unmatched is straightforward, as is having agents exit.

Second, our basic model also has no place for triads, which we previously saw to play a key role. This can be added directly, simply by having triples meet if they are all present in the town square. The extension is straightforward and thus omitted (see Chandrasekhar and Jackson (2018) for more detail on such an extension). In such a model, analogous to the pairs case above, it would be direct that LL and LLL could decline more than their counterparts (HH and LH; HHH, LHH, and LLH, respectively).

Third, note that it is plausible that when one aspect of a relationship becomes less important, there is some risk that the entire relationship breaks up, since there is a fixed costs of maintaining a relationship. By adding other types of links that are maintained and formed at the same time as financial links, the model can generate similar effects on other links as well. As we saw above in Table 6, when we look at advice-based links, the effects are more or less of the same magnitude in proportional terms and in the same direction as the financial links.

4.8. **Alternative Explanations.** In this section we try to address two issues. First, can we account for the facts without going to a model with undirected search while maintaining our assumptions about changes in payoffs? Second, are there alternative assumptions about changes in payoffs that can help us account for the facts in combination with a simpler model of network formation?

---

[33]See the calculations in Appendix J.

4.8.1. *Alternative models of network formation.* In Appendix D, we discuss four other models of network formation, variations of which are already in the literature. We show why one needs a model that goes beyond those models to generate the patterns in our data.

As such, our work contributes to the literature on network formation by introducing a model that combines features of different existing models, and showing why that combination of features is needed to match what we observe in the data.

Previous models of network formation that involve explicit choice by agents[34] have three flavors:

(i) models in which people have the opportunity to connect with whomever they want, subject to reciprocation (e.g., Jackson and Wolinsky (1996); Dutta and Mutuswami (1997); Bala and Goyal (2000); Currarini and Morelli (2000); Jackson and Van den Nouweland (2005); Herings, Mauleon, and Vannetelbosch (2009); Jackson, Rodriguez-Barraquer, and Tan (2012); Boucher (2015));

(ii) models in which there are exogenously random meetings and then, conditional upon meeting, people choose with whom to connect (e.g., Watts (2001); Jackson and Watts (2002); Christakis, Fowler, Imbens, and Kalyanaraman (2010a); König, Tessone, and Zenou (2014); Mele (2017)); and

(iii) models in which people put in some effort to socialize, which then results in some random meetings, but then relationships are formed as a result of those efforts without further choice (e.g., Currarini, Jackson, and Pin (2009, 2010); Cabrales, Calvó-Armengol, and Zenou (2011); Canen, Jackson, and Trebbi (2017)).

First, the empirical patterns that we observe here require a model with some externalities in the efforts to search and meet, which are not present in (i) or (ii). The basic point is that any model that gives agents full control over who they try to link with, which is what fails with undirected search, cannot explain why LL and LLL relationships drop in response to a decrease in Hs' willingness to link to Ls.[35] Our model relies on the idea that individuals put effort into trying to meet but cannot choose who they are meeting. The reason Hs lower effort less than Ls is that they are (correctly) more optimistic about actually linking with those who they meet. If on the other hand, the Ls could costlessly meet with each other, or they meet people at random at no cost and can decide who to pair with, LL and LLL links should go up.

Second, models in class (iii) allow for search efforts but then do not involve the choices of whom to connect to, as are present in (i) and (ii). This choice of whom to connect to is important in adjusting the incentives and the relative rates at which HL links form compared

---

[34]There is also a large literature of network formation that involves no strategic choice but just a stochastic model of network formation/evolution (e.g., see Jackson (2008) for some description and references). Those models are not equipped to match the data here.

[35]Recall that the base model does not include "friend-of-friend" effects. While models with such local spillover effects can indeed generate the fact that LL and LLL relationships drop, they cannot explain why LL and LLL relationships drop *more* than LH and LLH, as we discuss in Appendix D.3.

to HH or LL links, which is important in our analysis.[36] Thus, the model that we introduce is a hybrid of these three classes: effort is needed to meet others and affects the relative rates at which people are randomly met, but then choice is involved conditional upon meeting.

Thid, in addition, our model has two other features that help us to match the data. One is that effort is not only needed to meet new people, but also to maintain existing relationships - as the patterns we observe in the data exhibit similarities both in terms of which relationships are retained and which new ones are formed. The second is that socializing affects the opportunities to form multiple types of relationships at the same time - and so relationships are naturally "multiplexed."

The combination of all four of these features – efforts to socialize with rates of meetings dependent on relative efforts, mutual choice required to form relationships conditional upon meeting, effort needed to maintain relationships, and multiple types of relationships formed at the same time – allows us to capture all of the nuances and rich patterns that we observe in the data. In Appendix D, we discuss why dropping any one of these features would fail to capture some aspects of the data.

4.8.2. *Alternative models of match value.* We have so far assumed that match value depends only on types and does not depend on the pattern of matching. It is possible, for example, that matches are substitutes, so that when many LH links break, the value of LL links may go up. This would predict an increase in LL links, which goes in the wrong direction. It is also possible, though perhaps less likely, that links are complements: perhaps when an L can no longer borrow from the Hs, she gives up the entire project and therefore also stops borrowing from other Ls. However in this case the LL links break because some LH links have disappeared and therefore the effect on LL links will be smaller than the effect on LH links in proportional terms.

A similar possibility is that an L might want to link with another L because that second L is in turn linked to an H, and this is valuable for another reason (e.g., risk sharing, information exchange, network support, etc.) But, as above, in this case the LL links are breaking because LH links have disappeared and so again, the effect on LL links would be smaller than the effect on LH links in proportional terms. See Appendix D.3.2 for more details.

Another possibility is that the reason LL links drop is that Ls recognize that even if they don't participate in microfinance, it is available to them. This is probably true for some of them, but because we use microfinance eligibility to determine who is an L, it is less true for them than for the H (who also don't all borrow). An H is therefore more likely to break their link with an L on these grounds than another L.

---

[36]The models by Currarini, Jackson, and Pin (2009, 2010); Canen, Jackson, and Trebbi (2017) adjust the cross-type meeting rates either by a congestion meeting technology or a homophily parameter. Ours is derived from utility considerations, which helps understand why things change as microfinance is introduced.

Yet another alternative is based on the idea that the very fact that Hs tend to socialize with Hs in microfinance meetings would provide a force unique to participants, hence Hs, to form new links. This might crowd out their other links, but that would predict that LH links should decline by more than LL links (which should not be affected). We do not find this. We further examine this alternative in Online Appendix H. We show that our main results hold even if we condition on all pairs where neither member joined microfinance (86% of baseline links). Even under this restriction, the H classification has content – Hs are more suitable for microfinance by construction, and therefore even non-borrowers have higher option value from future access to microfinance.[37] However, we acknowledge that these results are only suggestive given that actual microfinance take-up is endogenous.

Another possibility is a slight variant of our undirected search model where the H types simply do not have time to meet with the Ls anymore. Notice our general form of the model allows for this.

A final possibility is that the entry of microfinance leads to rapid economic growth in the village, so that both H and L type don't need to maintain informal relationships any more. This is not only inconsistent with the extensive literature on microfinance, which finds little impact of microfinance entry on average village or neighborhood level outcomes such as consumption, investment or business profit (Angelucci, Karlan, and Zinman, 2015; Attanasio, Augsburg, De Haas, Fitzsimons, and Harmgart, 2015; Augsburg, De Haas, Harmgart, and Meghir, 2015; Banerjee, Karlan, and Zinman, 2015b; Banerjee, Duflo, Glennerster, and Kinnan, 2015a; Crépon, Devoto, Duflo, and Parienté, 2015; Tarozzi, Desai, and Johnson, 2015) (see Meager (2015) for a meta-analysis), but it is also inconsistent with the household finances and borrowing that we see going on in the villages.[38] Moreover, per se, this explanation would also not account for larger effects for the L households.

5. Informal Credit and Insurance

5.1. **Impact on Borrowing Patterns.** Next we examine how both microfinance and informal borrowing respond to the arrival of microfinance. In both the Karnataka and Hyderabad data we have rich borrowing information, and we can measure the impacts of microfinance on several different types of household borrowing (e.g., microfinance, friend, self-help group

---

[37]This option value may in turn reduce the value non-borrowing Hs receive from maintaining and forming network relationships.

[38]All of these studies, other than Augsburg et al. (2015), estimate the treatment effect at the community (urban neighborhood or village) level; these are therefore directly comparable to the effect on the village networks in our study. Moreover only Crépon et al. (2015) allows for the estimation of spillovers and does not find any support for the view that there are some people who get large benefits, while other lose out. One exception to this evidence of a lack of impact is Breza and Kinnan (2018) who find negative effects of shutting down microcredit in Andhra Pradesh, India. Their interpretation of this result is that there are spillovers, possibly from a demand shortage generated by a large-scale demand crunch. However their unit of comparison is a district, which has many hundreds of villages, and therefore there is a much bigger scope for spillovers. The villages in our paper are relatively far from each other and there is much less chance of cross-village spillovers.

member, family member, or money lender). If the loss in network links corresponds to a drop in informal financial transactions, then informal borrowing should respond in a manner similar to our network results, above.

We have data on the amount borrowed by source for the entirety of our sample. We begin by regressing the amount borrowed on dummies for microfinance village, post, and household type in the Karnataka sample:

$$y_{ivt} = \alpha + \beta_1 \text{MF}_v \times \text{Post}_t + \gamma_1 \text{MF}_v \times H_{iv} \times \text{Post}_t + \gamma_2 H_{iv} \times \text{Post}_t + \gamma_3 \text{MF}_v \times H_{iv}$$
$$+ \delta_1 \text{MF}_v + \delta_2 H_{iv} + \delta_3 \text{Post}_t + \delta' X_{ij,v} + \epsilon_{ivt},$$

where again y_{ivt} is the amount borrowed from the stated source (MFI, friends, self-help group, family, moneylenders).

Table 9, Panel A presents the results for the Karnataka data. In column 1, we find that L households do borrow from MFI (the classification is not perfect), the coefficient is 477 rupees ($p \leq 0.001$). However H borrow much more than L's do (1,787 rupees more). Column 2 finds that that L households lose more from their informal network than they gain from microfinance: they lose Rs. 1045 ($p = 0.083$) in loans from their entire network (friends, self-help groups, and family) after their village is exposed to microfinance. The H households lose additionally more, but this corresponds to their additional borrowing from MFIs.

Panel B of Table 9 turns to the Hyderabad data. Here we present impacts on borrowing measured in the first post-microfinance survey wave, 15-18 months after microfinance was introduced in treatment neighborhoods.[39]

$$y_{ivt} = \alpha + \beta_1 \text{MF}_v + \gamma_1 \text{MF}_v \times H_{iv} + \delta_2 H_{iv} + \delta' X_{ij,v} + \epsilon_{ivt},$$

Starting with column 1, we find no impact of MF on borrowing for L type in this case but a large and significantly greater impact for H types (8,776 rupees, $p \leq 0.001$).

We again observe a decline in informal borrowing for L types. All informal borrowing declines by Rs.4,341 ($p = 0.078$), relative to a control mean of Rs.36,347. Here, this decline is almost entirely offset for the H types, though the difference between the two is not statistically significant ($p = 0.137$).

Taken together, the evidence suggests that exposure to microfinance has an adverse effect on the network borrowing of the Ls that is large enough to dominate their small (if any) potential gains from microfinance. This is especially striking because, all else being the same, we would have expected their borrowing from friends to go down less than that of the Hs or even to go up to the extent there is re-lending (from Hs to Ls).

---

[39]Recall that we only measured networks in the third wave, four years later. However, we collected information about informal borrowing in the earlier waves.

5.2. **Impact on Risk Sharing.** One important role of village networks is risk sharing, both through gifts and through "soft" loans whose terms are state-contingent (see, e.g., Townsend (1994); Udry (1994)). Moreover, new policies can spill-over onto others through risk sharing relationships.[40] In light of our finding that access to formal credit results in the loss of network links and a reduction in informal borrowing, it is natural to investigate whether risk sharing is also reduced. If a household's ability to buffer income fluctuations through informal risk sharing is worsened when their village has access to microfinance, the loss of network links could have a direct negative impact on welfare. Microfinance practitioners often argue that even if borrowing households do not become richer because of the loans, they should at least be better able to smooth risk through continued access to credit. However, non-borrowing L households may see a worseining of risk sharing and consumption smoothing as they have fewer links without any direct benefits from microfinance.

Fortunately, the Hyderabad data contains panel information on both income and consumption[41], which allows us to run standard omnibus tests of risk sharing.

The loss of network degree is greater for the L types than the H types. In contrast, any direct benefits from microfinance borrowing are, by construction, lower for the L types relative to the H types. Thus, we expect that risk sharing should worsen for L households in treatment, compared to L households in control. The predictions are less clear for the H types, who may be able to smooth income risk using microfinance.

To test this prediction, we estimate a version of the standard regression from Townsend (1994), which allows the pass-through of income to consumption to differ by treatment status, differentially for H and L types. The subscript i indexes households, v indexes villages, and t indexes time. Microfinance (i.e., treated) villages are denoted with MF. The term α_i is a household-fixed effect which, as well as controlling for the household's time-invariant Pareto weight in a risk-sharing regression, absorbs the main effect of treatment status and of type (H or L). The term γ_{vt} is a area-time fixed effect which captures the aggregate shock to a particular neighborhood in a given survey wave. We estimate the following "long differences" specification, using the first and third waves of the data.[42]

---

[40]Angelucci and De Giorgi (2009), for example, analyze the Progresa cash transfer program in Mexico and find that even non-eligible households consume more. Moreover, Angelucci et al. (2018) trace these impacts through the kinship network. Albarran and Attanasio (2003) also analyze the interplay between policies and risk sharing, highlighting the potential for risk sharing to go down due to improved outside options.

[41]Recall, the Hyderabad data only contain network information collected in 2012, but income and consumption were collected in 2007-8, 2010 and 2012.

[42]As noted by Hayashi et al. (1996), long differences may perform better than one-period differences if income changes are dominated by transitory measurement error or transitory wage changes, or if income is known one period in advance. Moreover, in our setting, the second wave of data was collected at a point in time when both treatment and control areas had access to microfinance, which may limit our ability to detect treatment effects on risk-sharing. Results using all three waves of data are qualitatively similar but less precise, consistent with the measurement error concern noted by Hayashi et al. (1996).

$$c_{ivt} = \alpha_i + \gamma_{vt} + \beta_1 y_{ivt} + \beta_2 y_{ivt} \times MF_v + \beta_3 y_{ivt} \times H_i + \beta_4 y_{ivt} \times MF_v \times H_i + \delta' X_{ij,v} + \epsilon_{ivt}$$

The coefficient β_1 measures the extent of income pass-though for L households in control areas. The prediction that risk sharing should worsen for L households in treatment, compared to L households in control, is tested by the coefficient β_2. The coefficient β_3 captures the extent of income pass-though for H households in control areas. Finally β_4 measures any differential treatment effect on risk sharing for H households.[43]

Table 10 presents the results. In column 1, we consider total per capita household expenditure. In column 2, we consider non-food per capita consumption, which is typically more discretionary and therefore more responsive to shocks. We find that L households in control areas experience an INR 0.059 drop in non-food consumption for a INR 1 drop in income, a 5.9% pass-through rate ($p = 0.005$). However, for L households in treated areas, the pass-though increases by 0.080 ($p = 0.018$). Thus, the loss of network links for L households is economically consequential in terms of worsened ability to buffer income fluctuations. The estimates are qualitatively similar for total consumption, but slightly less precise ($p = 0.098$ for the test that pass through is greater in MF areas for L households).

The remaining coefficients in Table 10 consider how these patterns differ for H types. We find that microfinance causes the pass-through of income into non-food consumption for H types to decrease by 0.107 rupees, relative to the effect for L types ($p = 0.075$). For these types, microfinance causes no detectable change to income pass-through ($\beta_2 + \beta_4$, $p = 0.546$).

In summary, these results demonstrate that, while H households' ability to buffer income risk is unaffected by microfinance exposure, this is not true for L households. For them, the loss of network links arising from microfinance exposure has negative consequences: their ability to smooth income fluctuations is worsened markedly. Given the lack of microfinance's effect on income, these results show that the reduction in links for L types is, in all likelihood, welfare reducing.

6. Conclusion

By studying the introduction of microcredit to just a subset of communities in two different settings, we established that not only did the social networks change in response, but those who were least likely to take up microcredit experience substantial losses in links, even in groups (triangles) where no one was involved in microcredit. This is accompanied by a loss in the ability to borrow from informal networks for those households and a reduction in their ability to smooth risk (where we can measure it). The results were first obtained in one setting and then confirmed in a second, independent one.

---

[43]In Appendix Table L.1, we show that microfinance does not have any impact on income for either L or H types.

To explain the data, we developed a model in which agents put in effort in order to socialize, whom they meet has an undirected component, and agents engage in mutual consent to build links. Such a model features a global externality, beyond the typical externalities directly embedded in payoffs. When access to microfinance reduces a borrower's desire to maintain and form links with others, even those unlikely to join may reduce their own effort to maintain and build links. This is for two reasons: first, these low types who are linked to potential borrowers consequently have lower returns to such links (decreased payoff externality); second, because of the reduction in overall linking effort, even relationships where the direct payoffs are unaffected by microfinance can be affected. In equilibrium, those who are unlikely to be involved with microcredit may end up with the greatest losses in links.

The fact that our model provides patterns consistent with the data, of course, does not imply that it is the right or only mechanism behind the empirical observations. It will take further research to develop a full understanding of the forces underlying our empirical observations. Nonetheless, the facts – in particular the evidence of negative spillovers on the non-beneficiaries – have wide-ranging and important implications. The previous literature has shown that there may be important benefits from microfinance on participant households beyond the loans themselves, especially in terms of strengthened network connections. But if this comes at a significant cost of weakened connections in the rest of the community, this could worsen the aggregate ability of a community to buffer risk.

Regardless of the explanation for the changes, the more general lesson these findings illustrate is that social networks can involve spillovers, externalities, and complex relationships so that changing one part of the network can have quite extensive and unanticipated consequences. As a result, interventions into a community can change the social structure and interactions in ways that no one intended, with potentially large costs for some non-participants. Being mindful of these possibilities is important in designing effective policies.

REFERENCES

ALBARRAN, P. AND O. P. ATTANASIO (2003): "Limited commitment and crowding out of private transfers: evidence from a randomised experiment," *The Economic Journal*, 113, C77–C85.

ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of political economy*, 113, 151–184.

AMBRUS, A., M. MOBIUS, AND A. SZEIDL (2014): "Consumption risk-sharing in social networks," *American Economic Review*, 104, 149–82.

ANGELUCCI, M. AND G. DE GIORGI (2009): "Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption?" *American economic review*, 99, 486–508.

ANGELUCCI, M., G. DE GIORGI, AND I. RASUL (2018): "Consumption and investment in resource pooling family networks," *The Economic Journal*, 128, 2613–2651.

ANGELUCCI, M., D. KARLAN, AND J. ZINMAN (2015): "Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco," *American Economic Journal: Applied Economics*, 7, 151–82.

ARROW, K. J. (2000): "Observations on social capital," *Social capital: A multifaceted perspective (World Bank Publisher)*, 6, 3–5.

ATTANASIO, O., B. AUGSBURG, R. DE HAAS, E. FITZSIMONS, AND H. HARMGART (2015): "The impacts of microfinance: Evidence from joint-liability lending in Mongolia," *American Economic Journal: Applied Economics*, 7, 90–122.

AUGSBURG, B., R. DE HAAS, H. HARMGART, AND C. MEGHIR (2015): "The impacts of microcredit: Evidence from Bosnia and Herzegovina," *American Economic Journal: Applied Economics*, 7, 183–203.

BALA, V. AND S. GOYAL (2000): "A noncooperative model of network formation," *Econometrica*, 68, 1181–1229.

BANERJEE, A., E. BREZA, E. DUFLO, AND C. KINNAN (2019a): "Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs?" Tech. rep., National Bureau of Economic Research.

BANERJEE, A., A. CHANDRASEKHAR, E. DUFLO, AND M. O. JACKSON (2013): "Diffusion of Microfinance," *Science*, 341, DOI: 10.1126/science.1236498, July 26 2013.

BANERJEE, A., A. G. CHANDRASEKHAR, E. DUFLO, AND M. O. JACKSON (2019b): "Using gossips to spread information: Theory and evidence from two randomized controlled trials," *The Review of Economic Studies*, 86, 2453–2490.

BANERJEE, A., E. DUFLO, R. GLENNERSTER, AND C. KINNAN (2015a): "The miracle of microfinance? Evidence from a randomized evaluation," *American Economic Journal: Applied Economics*, 7, 22–53.

BANERJEE, A., D. KARLAN, AND J. ZINMAN (2015b): "Six randomized evaluations of microcredit: Introduction and further steps," *American Economic Journal: Applied Economics*, 7, 1–21.

BEAMAN, L. AND J. MAGRUDER (2012): "Who gets the job referral? Evidence from a social networks experiment," *American Economic Review*, 102, 3574–93.

BELLONI, A. AND V. CHERNOZHUKOV (2009): "Least squares after model selection in high-dimensional sparse models," *MIT Department of Economics Working Paper*.

BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014a): "High-dimensional methods and inference on structural and treatment effects," *The Journal of Economic Perspectives*, 28, 29–50.

——— (2014b): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81, 608–650.

Binzel, C., E. Field, and R. Pande (2013): "Does the Arrival of a Formal Financial Institution Alter Informal Sharing Arrangements? Experimental Evidence from Village India," Working paper.

Bloch, F., G. Genicot, and D. Ray (2008): "Informal insurance in social networks," *Journal of Economic Theory*, 143, 36–58.

Blumenstock, J. and X. Tan (2016): "Social Networks and Migration: Theory and Evidence from Rwanda," *Working Paper.*

Blumenstock, J. E., N. Eagle, and M. Fafchamps (2016): "Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters," *Journal of Development Economics*, 120, 157–181.

Boucher, V. (2015): "Structural homophily," *International Economic Review*, 56, 235–264.

Breza, E. (2016): "Field experiments, social networks, and development," *The Oxford handbook of the economics of networks.*

Breza, E. and A. G. Chandrasekhar (2019): "Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment," *Econometrica*, 87, 175–216.

Breza, E., A. G. Chandrasekhar, T. H. McCormick, and M. Pan (2019): "Consistently estimating graph statistics using Aggregated Relational Data," *arXiv preprint arXiv:1908.09881.*

——— (2020): "Using aggregated relational data to feasibly identify network structure without network data," *American Economic Review*, 110, 2454–84.

Breza, E. and C. Kinnan (2018): "Measuring the equilibrium impacts of credit: Evidence from the Indian microfinance crisis," Tech. rep., National Bureau of Economic Research.

Cabrales, A., A. Calvó-Armengol, and Y. Zenou (2011): "Social interactions and spillovers," *Games and Economic Behavior*, 72, 339–360.

Canen, N., M. O. Jackson, and F. Trebbi (2017): "Endogenous Networks and Legislative Activity," *SSRN http://ssrn.com/abstract=2823338.*

Chandrasekhar, A. and M. O. Jackson (2018): "A Network Formation Model Based on Subgraphs," *SSRN Working Paper: http://ssrn.com/abstract=2660381.*

Chandrasekhar, A. and R. Lewis (2014): "Econometrics of sampled networks," MIT working paper.

Christakis, N., J. Fowler, G. Imbens, and K. Kalyanaraman (2010a): "An Empirical Model for Strategic Network Formation," *NBER Working Paper.*

Christakis, N. A., J. H. Fowler, G. W. Imbens, and K. Kalyanaraman (2010b): "An empirical model for strategic network formation," Tech. rep., National Bureau of Economic Research.

Comola, M. and S. Prina (2014): "Do Interventions Change the Network? A Dynamic Peer Effect Model Accounting for Network Changes," Working paper.

CRÉPON, B., F. DEVOTO, E. DUFLO, AND W. PARIENTÉ (2015): "Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco," *American Economic Journal: Applied Economics*, 7, 123–50.

CURRARINI, S., M. O. JACKSON, AND P. PIN (2009): "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica*, 77, 1003–1045.

——— (2010): "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proceedings of the National Academy of Sciences*, 107, 4857–4861.

CURRARINI, S. AND M. MORELLI (2000): "Network formation with sequential demands," *Review of Economic Design*, 5, 229 – 250.

DUTTA, B. AND S. MUTUSWAMI (1997): "Stable networks," *Journal of economic theory*, 76, 322–344.

FAFCHAMPS, M. AND S. LUND (2003): "Risk-sharing networks in rural Philippines," *Journal of development Economics*, 71, 261–287.

FEIGENBERG, B., E. FIELD, AND R. PANDE (2013): "The economic returns to social interaction: Experimental evidence from microfinance," *The Review of Economic Studies*, rdt016.

FIELD, E., R. PANDE, J. PAPP, AND Y. J. PARK (2012): "Repayment flexibility can reduce financial stress: a randomized control trial with microfinance clients in India," *PloS one*, 7, e45679.

HAYASHI, F., J. ALTONJI, AND L. KOTLIKOFF (1996): "Risk-Sharing between and within Families," *Econometrica: Journal of the Econometric Society*, 261–294.

HERINGS, P. J.-J., A. MAULEON, AND V. VANNETELBOSCH (2009): "Farsightedly stable networks," *Games and Economic Behavior*, 67, 526–541.

HESS, S. H., D. JAIMOVICH, AND M. SCHÜNDELN (2020): "Development Projects and Economic Networks: Lessons From Rural Gambia," *The Review of Economic Studies.*

HOFF, P., A. RAFTERY, AND M. HANDCOCK (2002): "Latent Space Approaches to Social Network Analysis," *Journal of the American Statistical Association*, 97:460, 1090–1098.

JACKSON, M. O. (2003): "The stability and efficiency of economic and social networks," *Advances in Economic Design, Heidelberg: Springer-Verlag, edited by Koray, S. and Sertel, M.*

——— (2008): *Social and economic networks*, Princeton: Princeton University Press.

JACKSON, M. O., T. R. RODRIGUEZ-BARRAQUER, AND X. TAN (2012): "Social Capital and Social Quilts: Network Patterns of Favor Exchange," *American Economic Review*, 102, 1857–1897.

JACKSON, M. O. AND A. VAN DEN NOUWELAND (2005): "Strongly stable networks," *Games and Economic Behavior*, 51, 420–444.

JACKSON, M. O. AND A. WATTS (2002): "The evolution of social and economic networks," *Journal of Economic Theory*, 106(2), 265–295.

Jackson, M. O. and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.

Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl (2009): "Trust and Social Collateral," *The Quarterly Journal of Economics*, 24, 1307–1361.

Kinnan, C. and R. Townsend (2012): "Kinship and financial networks, formal financial access, and risk reduction," *American Economic Review*, 102, 289–93.

König, M., C. Tessone, and Y. Zenou (2014): "Nestedness in networks: A theoretical model and some applications," *Theoretical Economics, forthcoming.*

Meager, R. (2015): "Understanding the impact of microcredit expansions: A bayesian hierarchical analysis of 7 randomised experiments," *arXiv preprint arXiv:1506.06669.*

Mele, A. (2017): "A structural model of Dense Network Formation," *Econometrica*, 85, 825–850.

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

Putnam, R. (2000): *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster.

Tarozzi, A., J. Desai, and K. Johnson (2015): "The impacts of microcredit: Evidence from Ethiopia," *American Economic Journal: Applied Economics*, 7, 54–89.

Townsend, R. M. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591.

Udry, C. (1994): "Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria," *The Review of Economic Studies*, 61, 495–526.

Van Zandt, T. and X. Vives (2007): "Monotone equilibria in Bayesian games of strategic complementarities," *Journal of Economic Theory*, 134, 339–360.

Vera-Cossio, D. A. (2019): "Targeting Credit through Community Members," .

Watts, A. (2001): "A dynamic model of network formation," *Games and Economic Behavior*, 34, 331–341.

Tables

Table 1. Sample Statistics

Panel A: Karnataka Wave 1 Villages

	Obs	Control Mean	Control SD	Treatment - control		
				Coeff.	5% limit	p-value
Average Density	75	0.119	0.042	-0.020	0.015	0.011
Average Degree	75	13.355	2.443	1.472	1.148	0.014
Average Clustering	75	0.334	0.074	-0.041	0.030	0.009
Average Closeness	75	0.379	0.046	-0.013	0.019	0.183
Number of Households	75	165.812	48.945	57.397	24.354	0.000

Panel B: Karnataka Wave 1 Households

	Obs	Control Mean	Control SD	Treatment - control		
				Coeff.	5% limit	p-value
Eligible Female	7511	0.943	0.233	0.008	0.015	0.216
Harm mean distance to leaders	7511	0.474	0.079	-0.021	0.020	0.049
No access to latrine	7511	0.748	0.434	-0.038	0.051	0.205
Number of rooms	7511	2.489	1.313	-0.001	0.140	0.973
Thatched roof	7511	0.021	0.145	-0.002	0.014	0.678
Distance to Bangalore	7511	61.114	17.458	-3.823	8.074	0.309
All loans	7511	37861.564	129797.423	1351.740	11597.294	0.819
Network (friends and family) loans	7511	2735.470	25394.731	6.467	1716.401	0.994
SHG loans	7511	2543.994	6944.324	14.783	968.668	0.976
Bank loans	7511	19892.356	106358.225	3563.106	8808.589	0.428
Moneylender loans	7511	3638.339	20456.671	-164.660	1656.949	0.846
H	7511	0.542	0.498	-0.209	0.068	0.000

Panel C: Hyderabad

	Obs	Control Mean	Control SD	Treatment - control		
				Coeff.	5% limit	p-value
Total outstanding debt in area, baseline	105	39675.337	47776.778	-6981.245	13853.634	0.326
Area population, baseline	105	264.615	160.467	-3.385	58.378	0.910
Prime-aged (18-45) women in HH, endline 1	6865	1.456	0.820	-0.024	0.056	0.413
Owns land in Hyderabad, endline 1	6865	0.061	0.239	-0.001	0.014	0.897
Owns land in village, endline 1	6865	0.194	0.396	0.006	0.057	0.841
H	6865	0.241	0.427	-0.016	0.077	0.684

Notes: This table presents summary statistics and baseline balance from our sample from Karnataka (Panels A and B) and Hyderabad (Panel C). The network characteristics in Panel A are calulated for an unweighted, symmetric graph of households. A link is based on whether there is any of 12 relationship types between two nodes. The "5% limit" column shows how big the difference between treatment and control would have needed to be to be significant at the 5% level. The "p-value" column shows a test of significance on the difference between treatment and control. In Karnataka, there are 75 villages in our sample and 43 received microfinance. In Hyderabad, 102 villages were subject to randomized assignment of microfinance. In Panel C we exclude 15 villages where the number of households are topcoded, as these would bias the estimates of all these statistics.

TABLE 2. Characteristics of H versus L

Panel A: Karnataka - Demographics and Amenities variables

	(1) GMOBC	(2) Latrine	(3) Num. Rooms	(4) Num. Beds	(5) Thatched Roof	(6) RCC Roof
H	−0.221 (0.008) [0.000]	−0.115 (0.008) [0.000]	−0.177 (0.022) [0.000]	−0.213 (0.021) [0.000]	0.019 (0.003) [0.000]	−0.053 (0.006) [0.000]
Depvar Mean	0.7	0.261	2.36	0.84	0.0235	0.117
Observations	14,904	14,904	14,904	14,904	14,904	14,904

Panel B: Karnataka - Network variables

	(1) Degree	(2) Links to L	(3) Links to H	(4) Eig. Cent.
H	1.947 (0.124) [0.000]	−0.432 (0.085) [0.000]	2.355 (0.059) [0.000]	0.017 (0.001) [0.000]
Depvar Mean	8.97	4.61	3.09	0.0524
Observations	14,904	14,904	14,904	14,904

Panel C: Hyderabad - Demographics and Amenities variables

	(1) GMOBC	(2) Latrine	(3) Num. Rooms	(4) Thatched Roof	(5) RCC Roof
H	0.029 (0.048) [0.543]	0.066 (0.048) [0.176]	0.064 (0.151) [0.673]	0.005 (0.005) [0.358]	−0.014 (0.046) [0.764]
Observations	4,520	4,483	4,516	4,516	4,508
Depvar Mean	0.429	0.578	2.314	0.025	0.882

Panel D: Hyderabad - Network Variables

	(1) Exp. Degree	(2) Exp. Links to L	(3) Exp. Links to H	(4) Exp. Eig. Cent.
H	0.161 (0.240) [0.502]	−1.442 (0.336) [0.000]	1.601 (0.325) [0.000]	0.015 (0.007) [0.039]
Observations	4,519	4,519	4,519	4,523
Depvar Mean	5.813	4.353	1.463	0.075

Notes: Standard errors (clustered at the village level) are reported in parentheses. p-values are reported in brackets. Panels A and B pertains to Karnataka, based on Wave 1 data only. In Panels C and D pertaining to Hyderabad, the estimates reflect H-vs.L differences for the non-microfinance (control group) sample only.
GMOBC = A dummy for whether the household consists of general caste, otherwise backwards caste, so the omitted categories are scheduled caste and scheduled tribes. General and OBC are considered upper caste.

TABLE 3. Graph-Level Characteristics

Panel A: Karnataka

	(1) Density	(2) Density	(3) Density	(4) Clustering	(5) Clustering	(6) Clustering	(7) Closeness	(8) Closeness	(9) Closeness
Microfinance × Post	-0.0119	-0.0128	-0.0128	0.00357	0.00968	0.00968	-0.0225	-0.0153	-0.0155
	(0.00678)	(0.00690)	(0.00716)	(0.0146)	(0.0147)	(0.0153)	(0.00970)	(0.0117)	(0.0122)
	[0.0836]	[0.0669]	[0.0769]	[0.807]	[0.513]	[0.528]	[0.0234]	[0.193]	[0.208]
Microfinance	-0.0205	0.00477	0.00204	-0.0408	-0.0179	-0.00638	-0.0129	0.00947	0.00963
	(0.00842)	(0.00555)	(0.00227)	(0.0159)	(0.0148)	(0.00551)	(0.00993)	(0.0101)	(0.0106)
	[0.0175]	[0.393]	[0.373]	[0.0123]	[0.230]	[0.250]	[0.199]	[0.353]	[0.366]
Post	-0.0117	-0.0145	-0.0145	-0.00913	0.00852	0.00852	0.105	-0.0472	-0.00499
	(0.00576)	(0.0107)	(0.0111)	(0.0100)	(0.0249)	(0.0258)	(0.00762)	(0.0522)	(0.0778)
	[0.0454]	[0.182]	[0.198]	[0.366]	[0.733]	[0.742]	[0]	[0.369]	[0.949]
Observations	150	150	150	150	150	150	150	150	150
Double-Post LASSO		✓	✓		✓	✓		✓	✓
Village FE			✓			✓			✓
Non MF Mean	0.1135	0.1135	0.1135	0.329	0.329	0.329	0.431	0.431	0.431
Depvar Mean	0.0983	0.0983	0.0983	0.307	0.307	0.307	0.418	0.418	0.418

Panel B: Hyderabad

	(1) Density	(2) Density	(3) Clustering	(4) Clustering	(5) Closeness	(6) Closeness
Microfinance	−0.006	−0.007	−0.004	−0.004	−0.002	−0.002
	(0.003)	(0.003)	(0.007)	(0.007)	(0.002)	(0.002)
	[0.086]	[0.048]	[0.632]	[0.632]	[0.473]	[0.473]
Double-Post LASSO	No	Yes	No	Yes	No	Yes
Depvar Mean	0.0286	0.0286	0.0539	0.0539	0.00263	0.00263
Non MF Mean	0.0318	0.0318	0.0544	0.0544	0.00309	0.00309
Observations	4,429	4,429	89	89	89	89

Notes: Standard errors (clustered at the village level) are reported in parentheses. p-values are reported in brackets. For Panel A pertaining to Karnataka, controls consist of the share of upper caste households, number of households in the village, share of households in self-help groups, share Hindu, share with a latrine in the house, share that own the household, share that have electricity and share that are leaders. For Panel B pertaining to Hyderabad, controls are the same demographic characteristics of households and villages that are used in random forest classification of H vs L. Columns (1) and (2) for density were found through direct elicitation at the node level, whereas columns (3) - (6) were estimated at the graph level through ARD data. GBMOC is "General (caste), Backward (caste), Muslim, (or) Other Caste." RCC is Reinforced Cement Concrete.

Table 4. Link Evolution, Karnataka

	(1) Linked Post-MF	(2) Linked Post-MF	(3) Linked Post-MF	(4) Linked Post-MF
Microfinance	−0.058 (0.018) [0.002]	−0.060 (0.020) [0.003]	−0.023 (0.008) [0.006]	−0.021 (0.008) [0.007]
Microfinance × LH	0.009 (0.015) [0.573]	0.001 (0.014) [0.936]	0.007 (0.004) [0.120]	0.007 (0.004) [0.081]
Microfinance × HH	0.039 (0.022) [0.086]	0.023 (0.021) [0.280]	0.009 (0.007) [0.206]	0.013 (0.006) [0.040]
LH	−0.025 (0.012) [0.036]	−0.005 (0.011) [0.649]	−0.002 (0.004) [0.566]	−0.007 (0.004) [0.073]
HH	0.008 (0.017) [0.622]	0.041 (0.017) [0.017]	0.021 (0.006) [0.001]	0.008 (0.006) [0.193]
Observations	57,376	57,376	846,561	846,561
Linked Pre-MF	Yes	Yes	No	No
Controls		✓		✓
Depvar Mean	0.441	0.441	0.0636	0.0636
LL, Non-MF Mean	0.482	0.482	0.0753	0.0753
MF + MF x LH= 0 p-val	0.014	0.009	0.015	0.015
MF + MF x HH= 0 p-val	0.361	0.09	0.101	0.233
MF + LHxMF = MF + HHxMF p-val	0.137	0.275	0.641	0.231

Notes: Standard errors (clustered at the village level) are reported in parentheses. p-values are reported in brackets. Controls are selected by double post lasso among centrality controls (vector of flexible controls for centrality of both nodes), household characteristics (caste, a number of wealth proxies including number of rooms, number of beds, electrification, latrine presence, and roofing material) and all variables that are used in the random forest classification.

TABLE 5. Link Evolution, Hyderabad

	(1) Prob. Linked	(2) Prob. Linked
Microfinance	−0.006 (0.003) [0.023]	−0.006 (0.003) [0.035]
Microfinance × HH	−0.009 (0.009) [0.296]	−0.009 (0.008) [0.269]
Microfinance × LH	0.003 (0.003) [0.432]	0.002 (0.003) [0.470]
HH	0.019 (0.008) [0.021]	0.018 (0.008) [0.025]
LH	0.002 (0.003) [0.566]	0.002 (0.003) [0.571]
Observations	141,990	141,990
Controls	No	Yes
Depvar Mean	0.0255	0.0255
LL, Non MF Mean	0.0268	0.0268
MF + MF × HH = 0 p-val	0.097	0.081
MF + MF × LH = 0 p-val	0.458	0.396
MF + MF × HH = MF + MF x LH p-val	0.049	0.047

Notes: Standard errors (clustered at the village level) are reported in paren-theses. p-values are reported in brackets. The controls are selected by dou-ble post lasso among all the variables that are used for its random forest classification, and includes several household and village level characteris-tics.

TABLE 6. Link Evolution for Info and Financial Links

Panel A: Karnataka

	(1) Financial Linked Post-MF	(2) Financial Linked Post-MF	(3) Info Linked Post-MF	(4) Info Linked Post-MF
Microfinance	−0.052	−0.012	−0.050	−0.015
	(0.021)	(0.005)	(0.018)	(0.005)
	[0.014]	[0.016]	[0.006]	[0.007]
Microfinance $\times LH$	−0.005	0.003	0.002	0.006
	(0.019)	(0.003)	(0.017)	(0.003)
	[0.813]	[0.213]	[0.892]	[0.076]
Microfinance $\times HH$	0.029	0.004	0.040	0.005
	(0.026)	(0.005)	(0.021)	(0.005)
	[0.263]	[0.373]	[0.063]	[0.284]
LH	−0.015	−0.001	−0.014	−0.004
	(0.015)	(0.003)	(0.013)	(0.003)
	[0.331]	[0.623]	[0.288]	[0.183]
HH	−0.00002	0.013	−0.004	0.012
	(0.022)	(0.004)	(0.017)	(0.004)
	[1.000]	[0.003]	[0.813]	[0.005]
Observations	27,072	876,865	37,044	866,893
Linked Pre-MF	Yes	No	Yes	No
Depvar Mean	0.333	0.0341	0.326	0.0377
LL, Non-MF Mean	0.371	0.04	0.361	0.0464
MF + MF x LH = 0 p-val	0.005	0.035	0.008	0.025
MF + MF x HH = 0 p-val	0.343	0.157	0.615	0.061
MF + LHxMF = MF + HHxMF p-val	0.14	0.828	0.046	0.898

Panel B: Hyderabad

	(1) Financial Links	(2) Financial Links	(3) Non Financial Links	(4) Non Financial Links
Microfinance	−0.358	−0.414	−0.196	−0.164
	(0.138)	(0.136)	(0.118)	(0.116)
	[0.012]	[0.004]	[0.100]	[0.161]
Microfinance $\times H$	0.550	0.819	0.653	0.704
	(0.208)	(0.190)	(0.210)	(0.199)
	[0.010]	[0.000]	[0.003]	[0.001]
H	0.080	−0.363	−0.192	−0.387
	(0.144)	(0.152)	(0.164)	(0.168)
	[0.582]	[0.019]	[0.244]	[0.024]
Double-Post LASSO	No	Yes	No	Yes
Depvar Mean	4.24	4.24	2.87	2.87
MF + MF $\times H$ = 0 p-val	0.377	0.03	0.043	0.008
Observations	4,429	4,429	4,429	4,429

Notes: Standard errors (clustered at the village level) are reported in parentheses. *p*-values are reported in brackets.
Panel A: Columns 1-2 restrict to financial links, while columns 3-4 restrict to non-financial links. Columns 1 and 3 consider links that existed in Wave 1, while columns 2 and 4 consider pairs of nodes that were not linked in Wave 1.
Panel B: All columns include a full set of controls. Centrality controls are a vector of flexible controls (a polynomial) for centrality of both nodes. Household characteristics are caste and a number of wealth proxies including number of rooms, number of beds, electrification, latrine presence, and roofing material. Household predictor variables consist of all variables that are used in the random forest classification. In every case we include interactions of all of these network, demographic, and classification variables with microfinance.

TABLE 7. Triples Evolution, Karnataka

	(1) Full triangle linked Post-MF	(2) Full triangle linked Post-MF	(3) Any link in triangle survived Post-MF	(4) Any link in triangle survived Post-MF
Microfinance	−0.078	−0.069	−0.085	−0.081
	(0.029)	(0.026)	(0.023)	(0.019)
	[0.008]	[0.008]	[0.000]	[0.000]
Microfinance × LLH	0.026	0.014	0.043	0.034
	(0.021)	(0.019)	(0.018)	(0.015)
	[0.228]	[0.463]	[0.015]	[0.024]
Microfinance × LHH	0.054	0.026	0.057	0.039
	(0.030)	(0.024)	(0.025)	(0.018)
	[0.072]	[0.274]	[0.022]	[0.029]
Microfinance × HHH	0.093	0.045	0.087	0.058
	(0.042)	(0.036)	(0.031)	(0.026)
	[0.028]	[0.206]	[0.006]	[0.023]
LLH	−0.024	−0.002	−0.037	−0.023
	(0.018)	(0.017)	(0.014)	(0.013)
	[0.180]	[0.914]	[0.009]	[0.071]
LHH	−0.037	0.010	−0.032	−0.005
	(0.025)	(0.022)	(0.017)	(0.014)
	[0.133]	[0.636]	[0.053]	[0.720]
HHH	−0.025	0.045	−0.012	0.025
	(0.033)	(0.029)	(0.022)	(0.019)
	[0.454]	[0.123]	[0.593]	[0.197]
Observations	53,233	53,233	53,233	53,233
Linked Pre-MF	Yes	Yes	Yes	Yes
Controls		✓		✓
Depvar Mean	0.197	0.197	0.808	0.808
LLL, Non-MF Mean	0.252	0.252	0.864	0.864
MF + MF x HHH = 0 p-val	0.698	0.497	0.935	0.296
MF + MF x LLH= 0 p-val	0.023	0.03	0.022	0.019
MF + MF x LHH = 0 p-val	0.262	0.05	0.141	0.02
MF + MF x HHH = MF + MF x LLHp-val	0.076	0.359	0.093	0.338
MF + MF x HHH = MF + MF xLHH p-val	0.212	0.506	0.075	0.273
MF + MF x LLH= MF + MF x LHH p-val	0.122	0.475	0.409	0.738

Notes: Standard errors (clustered at the village level) are reported in parentheses. p-values are reported in brackets. Controls are selected by double post lasso among centrality controls (vector of flexible controls for centrality of both nodes), household characteristics (caste, a number of wealth proxies including number of rooms, number of beds, electrification, latrine presence, and roofing material) and all variables that are used in the random forest classification.

TABLE 8. Triples Evolution, Hyderabad

All variables x 1000	Full Triangle Linked	Full Triangle Linked
	(1)	(2)
Microfinance	−0.05	−0.05
	(0.03)	(0.03)
	[0.089]	[0.143]
Microfinance × LLH	0.04	0.04
	(0.03)	(0.03)
	[0.121]	[0.170]
Microfinance × LHH	0.00	0.01
	(0.1)	(0.1)
	[0.980]	[0.935]
Microfinance × HHH	−0.2	−0.2
	(0.1)	(0.1)
	[0.205]	[0.266]
LLH	−0.02	−0.02
	(0.03)	(0.03)
	[0.429]	[0.580]
LHH	0.1	0.1
	(0.1)	(0.1)
	[0.311]	[0.324]
HHH	0.3	0.3
	(0.1)	(0.1)
	[0.039]	[0.050]
Observations	3,341,002	3,341,002
Controls	No	Yes
Depvar Mean	6.82e-02	6.82e-02
LLL, Non-MF Mean	8.09e-02	8.09e-02
MF + MF × HHH = 0 p-val	0.106	0.144
MF + MF × LLH = 0 p-val	0.869	0.913
MF + MF × LHH = 0 p-val	0.476	0.54
MF + MF × HHH = MF + MF × LLH p-val	0.095	0.126
MF + MF × HHH = MF + MF × LHH p-val	0.072	0.09
MF + MF × LLH = MF + MF × LHH p-val	0.387	0.447

Notes: Standard errors (clustered at the village level) are reported in parentheses. p-values are reported in brackets. The controls are all the variables that are used for its random forest classification, and includes several household and village level characteristics.

TABLE 9. Borrowing patterns

Panel A: Borrowing Patterns, Karnataka

	(1) MFI	(2) Informal (Non-Family)	(3) Friends	(4) SHG	(5) Moneylender	(6) Family
Microfinance × Post	476.917 (97.965) [0.000]	−1,045.365 (601.974) [0.083]	−655.808 (170.690) [0.000]	−856.257 (178.289) [0.000]	701.628 (575.630) [0.223]	562.584 (499.776) [0.261]
Microfinance × Post × H	1,786.513 (153.972) [0.000]	−1,675.480 (940.653) [0.075]	224.293 (266.722) [0.401]	19.434 (280.220) [0.945]	−2,197.763 (904.724) [0.016]	−1,282.628 (785.503) [0.103]
Microfinance × H	−0.369 (108.137) [0.998]	363.865 (660.550) [0.582]	−75.549 (187.299) [0.687]	260.618 (196.803) [0.186]	201.500 (635.404) [0.752]	779.004 (551.673) [0.158]
Post × H	190.818 (119.758) [0.112]	1,358.352 (732.537) [0.064]	−439.807 (207.711) [0.035]	108.243 (217.951) [0.620]	1,830.107 (703.683) [0.010]	284.331 (610.954) [0.642]
Observations	28,062	27,194	27,194	28,062	28,062	28,062
Depvar Mean	592.86	2426.289	914.773	1888.637	1515.799	2655.801
L, Non-MF Mean	188.108	2939.154	1238.718	1950.323	1571.435	2329.516
MF x Post x H + MF x Post =0 p-val	0	0.048	0.23	0.115	0.358	0.085

Panel B: Borrowing Patterns, Hyderabad

	(1) MFI	(2) Informal (Non-Family)	(3) Friends	(4) SHG	(5) Moneylender	(6) Family
Microfinance	−338.369 (368.610) [0.359]	−4,341.590 (2,463.045) [0.078]	177.504 (1,034.104) [0.864]	−1,295.257 (921.507) [0.160]	−2,962.479 (1,790.754) [0.099]	−411.817 (713.689) [0.564]
Microfinance × H	8,772.677 (603.006) [0.000]	3,729.850 (4,256.022) [0.381]	−489.241 (1,830.265) [0.790]	−1,379.784 (1,612.587) [0.393]	4,833.537 (2,917.249) [0.098]	1,911.666 (1,470.142) [0.194]
H	−137.457 (384.012) [0.721]	−4,306.400 (3,093.509) [0.164]	−2,329.828 (1,370.129) [0.090]	621.727 (1,240.245) [0.617]	−1,989.545 (1,966.270) [0.312]	−604.691 (833.654) [0.469]
Observations	6,811	6,863	6,863	6,863	6,863	6,863
Depvar Mean	3138.51	35298.64	8331.35	7162.35	19331.47	2698.52
L, Non MF Mean	2105.41	36346.96	8490.13	7224.29	20113.67	2753.41
MF + MF × H = 0 p-val	0	0.886	0.845	0.106	0.51	0.336

Notes: These tables present the effect of microfinance access on the loan amounts borrowed from various sources. Panel A pertains to Karnataka (outcomes are winsorized to the 1% level) and tracks loan amounts from microfinance institutions, friends, family, banks and moneylenders. All of its columns control for surveyed in wave 1 fixed effects. The average rate of inflation over the period between waves in Karnataka was 8% meaning a total of 65%. Panel B pertains to Hyderabad (outcomes are winsorized to the 1% level). Outcomes are measured in the first survey wave (2007-08). Here all specifications include demographic household and village controls (the same ones used in random forest classification of H vs L) subject to double-post LASSO. Standard errors (clustered at the village level) are reported in parentheses. p-values are reported in brackets.
MFI: Microfinance Institution
SHG: Self-Help Group

Table 10. Risk sharing, Hyderabad

	(1) Expenditures: Total	(2) Expenditures: Non-Food
Household Income per capita	0.111	0.059
	(0.027)	(0.021)
	[0.000]	[0.005]
Microfinance $\times$ Income	0.069	0.080
	(0.041)	(0.034)
	[0.098]	[0.018]
Household Income per capita $\times H$	0.072	0.032
	(0.051)	(0.034)
	[0.157]	[0.351]
Microfinance $\times$ Income $\times H$	-0.121	-0.107
	(0.074)	(0.060)
	[0.103]	[0.075]
Observations	10452	10361
Depvar Mean	1192.937	1192.937
L, Non MF Depvar Mean	1187.527	1187.527
Income Mean	1436.279	1439.993
L, Non MF Income Mean	1434.170	1436.413
Test: MF $\times$ Income + MF $\times$ Income $\times H$	0.348	0.546

Notes: Income is total household, monthly per capita earnings from employment or business activities, excluding private and government transfers. Dependent variable is monthly per capita household expenditure. In col. 1, expenditure excludes food and in col. 2, we present total expenditure. Data is from the first (2007-08) and third (2012) waves of the Hyderabad survey. Regression includes controls for household fixed effects and wave-by-neighborhood-by-type fixed effects. Additional controls are selected by double post lasso from the set of variables used in the prediction exercise, interacted with type. Standard errors (clustered at the neighborhood level) are reported in parentheses. p-values are reported in brackets.

APPENDIX A. PROOF OF PROPOSITION 1

We show there is a unique equilibrium and characterize it, here letting each agent's utility be fully dependent upon their label i.

From our discussion above, it follows directly that a best response must satisfy[44]

$$e_i = \min\left\{1, \; \frac{1}{c_i}\left(u_i + \sum_{j \neq i} \mathrm{E}^+[v_{ij}]\left(1 - F(-v_{ij})\right)\left(1 - F(-v_{ij})\right) e_j\right)\right\}.$$

Given the bound that $e_j \leq 1$, and the fact that $u_i > 0$, it follows that for sufficiently large c_i,

$$e_i = \frac{1}{c_i}\left(u_i + \sum_{j \neq i} \mathrm{E}^+[v_{ij}]\left(1 - F(-v_{ij})\right)\left(1 - F(-v_{ij})\right) e_j\right),$$

and is strictly between 0 and 1.

Thus, taking c_i to be sufficiently large for each i, we let u be the n-dimensional vector with entries $\frac{1}{c_i}u_i$ and E be the $n \times n$ matrix with ij entries

$$\frac{1}{c_i}\mathrm{E}^+[v_{ij}]\left(1 - F(-v_{ij})\right)\left(1 - F(-v_{ij})\right)$$

Then the characterization of equilibria can be written as

$$e = u + Ee,$$

which has a (unique) solution of $e = (I - E)^{-1}u$, given that E has nonnegative values that are less than 1 and so $(I - E)$ is invertible.

Note that two agents of the same type take the same effort by the symmetry of the expected utility in type and uniqueness of equilibrium overall.

Rewriting u to be the $|\Theta|$-dimensional vector with entries $\frac{1}{c_\theta}u_\theta$ and E to be the $|\Theta| \times |\Theta|$ matrix with θ, θ' entries

$$\frac{1}{c_\theta}\mathrm{E}^+[v_{\theta\theta'}]n_{\theta\theta'}\left(1 - F(-v_{\theta\theta'})\right)\left(1 - F(-v_{\theta'\theta})\right)$$

the unique equilibrium is given by

$$e = (I - E)^{-1}u.$$

The result on the comparative statics follows from Proposition 16 in Van Zandt and Vives (2007), noting the strict monotonicity of the best responses in the payoffs and actions of others and the interiority of the equilibrium.

---

[44]This drops the $n_{\theta\theta'}$ terms, but one can include an indicator n_{ij} and nothing in the argument below changes.