Social Investments, Informal Risk Sharing and Inequality

Attila Ambrus‡ Arun G. Chandrasekhar‡ Matt Elliott§

September 2014

Abstract

This paper investigates costly social investments, in the context of risk-sharing. Extending the bargaining micro-foundations of Stole and Zwiebel (1996), we postulate that the benefits of risk-sharing are distributed according to the Myerson value. In particular, more centrally connected individuals receive a higher share of the surplus. Our main focus is comparing individual versus social incentives to establish relationships. If agents in the community are homogenous, there is never underinvestment relative to the socially efficient benchmark. In contrast, there can be severe overinvestment. We find a novel trade-off between efficiency and equality, and show that the most stable efficient network is also the most unequal one. When there are multiple groups in a society, and incomes are more correlated within groups, underinvestment across groups is possible and more central agents have better incentives to form across-group links. Using data from 75 Indian village networks, we provide empirical evidence consistent with our model’s predictions. We show that the comparative statics of network structure as economic environmental parameters are changed are consistent with our theory, using a difference-in-difference approach looking at the risk-sharing network versus the friendship network.

1 Introduction

In the context of missing formal insurance markets, and limited access to lending and borrowing, incomes may be smoothed through informal risk sharing arrangements that utilize social connections and mitigate utility losses. A large theoretical and empirical literature studies this, but less attention has been paid to the network formation problem and the social investments that enable risk sharing. We study whether social investments are efficient. If not, is too much time allocated to maintaining relationships or too little? Moreover, do the forces pushing towards under and/or over investment generate inequality in society even when agents are ex-ante homogenous? And if so, does social inequality translate into financial inequality?

*We thank the National Science Foundation, SES-1155302, for funding the research that collected the data we use. We thank Nageeb Ali, Ben Golub, Willemien Kets, Cynthia Kinnan, Rachel Kranton, Peter Landry and Horacio Larreguy for helpful comments.

†Department of Economics, Duke University; email: attila.ambrus@duke.edu
‡Department of Economics, Stanford University and NBER; email: arungc@stanford.edu
§Division of the Humanities and Social Sciences, California Institute of Technology; email: melliott@hss.caltech.edu

Both underinvestment and overinvestment in social capital are conceivable. Two people establishing a social connection to share risk gain access to a less stochastic income stream which might generate improved opportunities to share risk with their other connections. As these positive spillovers might not be fully taken into account when deciding whether to establish the link, underinvestment can prevail. On the other hand, if more socially connected individuals receive a higher share of the surplus generated by risk sharing, that can lead to overinvestment. Villagers may form links to redistribute the surplus towards themselves, rather than to increase the overall surplus generated. The empirical literature also suggests that both types of inefficiencies are possible, in different contexts. Austen-Smith and Fryer (2006) cites numerous references from sociology and anthropology, suggesting that members of poor communities allocate inefficiently large amounts of time to activities maintaining social ties, instead of productive activities. In contrast, Feigenberg et al. (2013) find evidence in a microfinance setting that it is relatively easy to experimentally intervene and create social ties among people that yield substantial benefits. One explanation for this finding is that there is underinvestment in social relationships.

It is important to study whether there is too little or too much investment into social relations, both to put related academic work (which often takes social connections to be exogenously given) into context, and to guide policy choices. Consider the example of microfinance. If there is overinvestment, microfinance has a greater scope for efficiency savings in terms of reducing people's allocation of time into social investments. With underinvestment, however, it has more scope for smoothing incomes. If there is neither under nor overinvestment, it also tells us that informal risk sharing is working relatively well as a second-best solution. Understanding which regime applies can help anticipate policy implications and evaluate welfare impacts of interventions.

To explore efficiency and inequality, in this paper we consider a two-stage model of network-formation and risk-sharing, in a context in which agents with constant absolute risk aversion (CARA) utilities face uncertain endowment realizations. In the first stage, agents choose with whom to form connections. Link formation is costly, as in Myerson (1991) and Jackson and Wolinsky (1996). In the second stage connected agents commit to a risk-sharing arrangement that is contingent on future endowment realizations. We show that in our CARA setting expected utilities are transferrable through state-independent transfers, and efficient risk-sharing arrangements on any network component are uniquely pinned down up to these transfers. The latter determine the allocation of surplus among agents. To keep the model tractable, we abstract away from the issues of how to enforce risk-sharing agreements.

We assume that the social surplus generated by efficient risk-sharing arrangements is distributed among the agents according to the Myerson value, a network-specific version of the Shapley value. Our motivation here comes from two sources. First, if the surplus division is chosen in a centralized manner, then it has normative appeal on the grounds of fairness: two agents benefit equally from a social relationship between them, and receive benefits proportional to their average contributions to total surplus (from establishing costly

---

2Although we consider a model in which there is perfect risk sharing of income, we could easily extend the model so that some income is perfectly observed, some income is private and there is perfect risk sharing of observable income and no risk sharing of unobservable income. This would be consistent with the theoretical predictions of Cole and Kocherlakota (2001) and the empirical findings of Kinnan (2011). In the CARA utilities setting such unobserved income outside the scope of the risk-sharing arrangement does not affect our results.

3See Ambrus et al. (2014) for an investigation of such issues.

Second, we show that a simple and natural decentralized procedure leads to the same outcome, providing microfoundations for the Myerson value in our setting. The procedure combines the exchange algorithm of Bramoulle and Kranton (2007a) and the pairwise robustness to renegotiation requirement of Stole and Zwiebel (1996). Following Bramoulle and Kranton (2007a), efficient risk sharing is obtained by assuming that after every realization of endowments, there is an infinite sequence of pairwise exchanges between neighbors in which joint resources are split equally. This process equalizes the consumptions of the agents on a connected component of the network. However, given this social norm, all connected agents receive the same consumption independent of the structure of the social network. This for example means that agents with more social connections have to pay higher costs towards maintaining these links, but receive no additional benefits from doing so. To address this issue, we allow neighbors ex ante to engage in bilateral bargaining over state-independent transfers. For this part, we extend the canonical bargaining framework of Stole and Zwiebel (1996) to apply to any network. In each pairwise negotiation, the two agents agree to a transfer that evenly splits the surplus generated by the link, relative to their expected utilities in the absence of the link. This can be thought of as the transfers being robust to renegotiation, if renegotiation would result in a ‘splitting the difference’ outcome. From this exercise we obtain a recursive definition of how surpluses get divided on different networks. By applying the axiomatization of the Myerson value provided in Myerson (1980), we show that the unique division of surplus compatible with this recursive definition is the Myerson value. In doing so, we provide new foundations for the Myerson value by extending a result from Stole and Zwiebel (1996). Moreover, we do so simply by applying an axiomatization provided by Myerson (1980).

A key implication of the Myerson value determining surpluses is that more centrally connected agents receive a higher share of the surplus. Moreover, in our risk-sharing context it implies that agents receive larger payoffs from providing ‘bridging’ links to otherwise socially distant agents, than from providing local connections. Empirical evidence supports this feature of our model—see Goyal and Vega-Redondo (2007), and references therein from the organizational literature: Burt (1992), Podolny and Baron (1997), Ahuja (2000), and Mehra et al. (2001).

In the network formation stage, we study the set of pairwise stable networks (Jackson and Wolinsky, 1996). Our general analysis considers a community comprised of different groups, where all agents are ex ante identical within groups, and establishing links within groups is cheaper

5These motivations make the Myerson value a commonly used concept in the network formation literature. See a related discussion on p 422-425 of Jackson (2010).

6Stole and Zwiebel (1996) model bargaining between many employees and an employer. This scenario can be represented by a star network with the employer at the center.

7The process is decentralized, as it involves pairwise renegotiations. The result we extend is Theorem 1 in Stole and Zweibul (1996). Their Theorem 2 can also be extended to our setting, and this would provide fully noncooperative foundations. Indeed, related noncooperative foundations are provided by Fontenay and Gans (2004), while Navarro and Perea (2013) take a different approach to microfounding the Myerson value. Slikker (2007) also provides noncooperative foundations, although the game analyzed is more centralized: offers are made at the coalitional level.

8More precisely, in Section 4 we introduce the concept of Myerson distance to capture the social distance between agents in the network, and show that a pair of agents’ payoffs from forming a relationship are increasing in this measure.

9Results from Calvo-Armengol and Ilkilic (2009) imply that under some parameter restrictions - for example when agents are ex ante identical - the set of pairwise stable outcomes is equivalent to the (in general more restrictive) set of pairwise Nash equilibrium outcomes.
than across groups. We also assume that the endowment realizations of agents within groups are more positively correlated than across groups. Groups can represent different ethnic groups or castes in a given village, or different villages. The core results show that there can only be overinvestment within groups but no underinvestment, whereas, across groups underinvestment is likely to be the main concern.

To see the intuition about overinvestment within groups, we first consider the case of homogenous agents, that is when there is only one group. Using the inclusion-exclusion principle from combinatorics, we provide a complete characterization of stable networks. We show that in this case there can never be underinvestment in social connections, as agents establishing an essential link (connecting two otherwise unconnected components of the network) always receive a benefit exactly equal to the social value of the link. However, overinvestment, in the form of redundant links, is possible, and becomes widespread as the cost of link formation decrease. We also find a trade-off between efficiency and equality. Among all possible efficient network structures, we find that the most stable (for the largest set of parameter values) is the star, which also results in the most unequal division of surplus. The intuition is that that the star network minimizes the incentives of peripheral agents to establish redundant links. Conversely, the least stable efficient network entails the most equal division of surplus among all stable networks. Although agents are ex-ante identical, efficiency considerations push the structure of social connections towards asymmetric outcomes that elevate certain individuals. Socially central individuals emerge endogenously from risk-sharing considerations alone.

Turning attention to the case of multiple groups, we find that across group underinvestment becomes an issue when the cost of maintaining links across groups is sufficiently high. The reason is that the agents who establish the first connection across groups receive less than the total surplus generated by the link, providing positive externalities for peers in their groups. This gap between private and social benefits is smaller for agents located more centrally in their own group, providing a second force for some agents within a group to be more central. For two groups, we show that the most stable efficient network structure involve “stars” within groups, connected by their centers, and we establish a weaker form of this result for more than two groups. This reinforces the trade-off between efficiency and equality, in the many groups context.

Using data from 75 Indian villages, we provide some supporting evidence for our model. We split the villagers into two groups, by caste. From the theoretical analysis, risk sharing links are most valuable when they bridge otherwise unconnected components. And when a link does not provide such a bridge, its value depends on how far apart, suitably defined, the agents would otherwise be on the social network. We call this distance between a pair of agents their Myerson distance. Our theory predicts that there is an upper bound on the Myerson distance between any two unconnected agents within the same group, beyond which the pair of agents would have a profitable deviation by forming a link. In addition we predict that there will be inequality in social positions and that more central agents within their group will form across group links. However, there are many alternative stories consistent with these

---

10This is subject to a regularity condition that it is socially valuable for two agents from the same caste, if they would otherwise be isolated, to form a link and share risk together.

11In certain settings, over time, these central individuals could also establish more formal financial institutions, creating more entrenched inequality.

12While across group overinvestment remains possible, when across group link costs are relatively high, the main concern is underinvestment.

13There is an extensive literature that examines caste as a main social unit where risk sharing takes place (Townsend (1994), Munshi and Rosenzweig (2009), and Mazzocco and Saini (2012)).
predictions. We therefore also generate more subtle predictions to test: how changes in the economic environment in terms of income variability or correlation correspond to changes in network structure under our model, while differencing out the friendship network from the risk-sharing network.

Consider two villagers from the same caste. As income variability increases, or within-caste incomes become less correlated, all else equal, the value of a risk-sharing link between these villagers increase; the Myerson distances that can be observed in a stable network therefore decreases, which in turn implies the following: (i) income variability is positively associated with lower Myerson distances between unconnected agents; and (ii) within-caste income correlation is associated with higher Myerson distances. The theory also predicts that villagers have to be sufficiently central within their own caste (the threshold depending again on income variability and within- versus across-caste income correlation), to be incentivized to provide a risk-sharing link across castes. This yields our final predictions: (iii) in villages with more income variability, more agents will have sufficient incentives to form an across-caste link and so the association between within-caste centrality and who provides across-caste links is weaker; and (iv) in villages with more within-caste income correlation relative to across caste income correlation, more agents will again have sufficient incentives to form across caste links and so the association between within-caste centrality and having an across-caste link will again be weaker. Because working with the exact Myerson distances is computationally infeasible, we develop an approximation which is exact for tree graphs, and also check that our results are robust to other notions of network sparsity. We demonstrate that our predictions are borne out in our data.

To strengthen our results, we exploit the fact that we have multigraph data. Not only do we have complete financial network data for every household in every village, but we have complete friendship network data as well. As our theory only pertains to the financial network, we are able to take a differences-in-differences approach. For example, for predictions (i) and (ii), we look within villages, across network type and ask whether the association with economic environmental parameters (income variability and within-caste income correlation) differentially vary with the Myerson distance of the financial network compared to that of the social network. This allows us to take out arbitrary village-level fixed effects, and we find that our results are robust to such an analysis.

Ultimately, our empirical approach allows us to be more conservative than similar studies in the literature (e.g., Karlan et al. (2007), Ambrus et al. (2014), Kinnan and Townsend (2014)). While the studies above have access to just a few of networks (e.g., two, one and 16, respectively), we have 75 networks and also have multigraph data. Most studies, therefore, are forced to do statistical inference within networks – which limits the amount of correlated shocks they are able to handle. Relative to this approach, our focus on the village level and differencing out the social network, is extremely conservative.

On the theory side, the studies on social networks and informal risk-sharing that are most related to ours include Bramoulle and Kranton (2007a,b), Bloch et al. (2008), Jackson et al. (2012), Billand et al. (2012), Ali and Miller (2013a,b) and Ambrus et al. (2014). Among these papers, Bramoulle and Kranton (2007a,b) and Billand et al. (2012) investigate costly network formation. Bramoulle and Kranton’s (2007a,b) model assumes that the surplus on a connected income component is equally distributed, independently of the network structure. This rules out the possibility of overinvestment, and leads to different types of stable networks than in our model. Instead of assuming optimal risk-sharing arrangements, Billand et al. (2012) assume an exogenously given social norm, which prescribes that high-income agents transfer a fixed amount of resources to all low-income neighbors. This again leads to very
different predictions regarding the types of networks that form in equilibrium.

More generally, network formation problems are important. Establishing and maintaining social connections (relationships) is costly, in terms of time and other resources. However, on top of direct consumption utility, such links can yield many economic benefits. Papers studying formation in different contexts include Jackson and Wolinsky (1996), Bala and Goyal (2000), Kranton and Minehart (2001), Hojman and Szeidl (2008), and Elliott (2013). Notable in this literature is a lack empirical work, which can be attributed to a number of innate difficulties that taking these models to data presents. One common problem is a multiplicity of stable networks. But perhaps most important, is that the networks in question can only very rarely be partitioned into a sizeable number of separate networks that can reasonably be treated as independent. Since predictions are often at the level of the overall network structure, this makes testing extremely challenging. Our many observations of social networks that are relatively independent of each other coupled with our approach to circumventing data limitations, allow us to provide a first step towards testing predictions based on the overall network structure. And although we study a quite specific network formation problem tailored to risk-sharing in villages, the general structure of our problem is relevant other applications.\(^{14}\)

The remainder of the paper is organized as follows. Section 2 describes risk sharing on a fixed network. In Section 3 we introduce a game of network formation with costly link formation. We focus on network formation within a single group first in Section 4 and then turn to the formation of across-group links in Section 5. We empirically test predictions of our model in the data in Section 6. Section 7 concludes.

\section{Preliminaries: Risk-Sharing on a Fixed Network}

We consider an economy in which agents face stochastic income realizations, but can insure against this uncertainty through redistributions made over the network of social connections. Ultimately we are interested in investigating endogenous network formation. However, in order to define a noncooperative game of investing into social connections, first we will specify the risk-sharing arrangement that prevails for any given network.

\textit{The social structure}

We denote the set of agents in our model by \( N \), and assume that they are partitioned into a set of groups \( M \). We let \( G : N \to M \) be a function that assigns each agent to a group; i.e., if \( G(i) = g \) then agent \( i \) is in group \( g \). One interpretation of the group partitioning is that \( N \) represents individuals in a village, and the groups correspond to different castes. Another possible interpretation is that \( N \) represents individuals in a larger geographic region (such as a district or subdistrict), and groups correspond to different villages in the region.

The social network is represented by an undirected graph \( L \) on the set of nodes corresponding to agents in \( N \) such that \( l_{ij} \in L \) is interpreted as a link exists between agents \( i \) and \( j \). The social network influences both the set of feasible risk-sharing arrangements and the distribution of surplus from risk sharing, as described below. There can be links both between agents in the same group and between agents in different groups.

\(^{14}\)For a different more specific and more removed application, suppose researches can collaborate on a project. Each researcher brings something heterogeneous and positive to the value of the collaboration so that the value of the collaboration is increasing in the set of agents involved. Collaboration is only possible when among agents who are directly connected to another collaborator, and surplus is split according to the Myerson value (as in our work, motivated by robustness to renegotiations). Such a setting fits into our framework.
We will refer to $N(i; L) \equiv \{j : l_{ij} \in L\} \subset N$ as agent $i$’s neighbors. An agent’s neighbors can be partitioned according to the groups they belong to. Let $N_g(i; L)$ be $i$’s neighbors on network $L$ from group $g$. We will sometimes refer to subsets of agents $S \subseteq N$ and denote the subgraphs they generated by $L(S) \equiv \{l_{ij} \in L : i, j \in S\}$. A subset of agents $S \subseteq N$ is path connected on $L$ if, for each $i \in S$ and each $j \in S$, there exists a sequence of agents (a path) $\{i, k, k', \ldots, k'', j\}$ such that every pair of adjacent agents in the sequence is linked. For any network there is a unique partition of $N$ such that there are no links between agents in different partitions but all agents within a partition are path connected. We refer to the cells of this partition as network components. The shortest path between two path connected agents is the path with the shortest sequence of agents. The diameter of a network component, $\text{d}(C)$, is the maximum length of the shortest path between any two agents in $C$. A network component is a tree when there is a unique path between any two agents in the component. The degree centrality of an agent is simply the number of neighbors he has (i.e. the cardinality of $N(i)$).

**Incomes and Consumption**

Agents in $N$ face uncertain income realizations. For tractability, we assume that incomes are jointly normally distributed, with expected value $\mu$ and variance $\sigma^2$ for each agent.\(^{15}\) We assume that the correlation coefficient between the incomes of any two agents within the same group is $\rho_w$, while between incomes of any two agents not in the same group is $\rho_a < \rho_w$.\(^{16}\) That is, we assume that incomes are more positively correlated within groups than across groups, so that all else equal, social connections across groups have a higher potential for risk-sharing.

Although we introduce the possibility of correlated incomes in a fairly stylized way, our paper is one of the first to permit differentially correlated incomes between different groups. Such correlations are central to the effectiveness of risk-sharing arrangements, as shown below.

We refer to possible realizations of the vector of incomes as states, and denote a generic state by $\omega$. We let $y_i(\omega)$ denote the income realization of agent $i$ in state $\omega$.

Agents can redistribute realized incomes, as described below; hence their consumption levels can differ from their realized incomes. We assume that all agents have constant absolute risk aversion (CARA) utility functions:

$$u(c_i) = -\frac{1}{\lambda} e^{-\lambda c_i},$$

where $c_i$ is agent $i$’s consumption and $\lambda > 0$ is the coefficient of absolute risk aversion.

**Efficient Risk-Sharing Agreements**

We assume that income can only be directly shared between agents $i, j \in N$ if they are connected, i.e. $l_{ij} \in L$. However, through a sequence of bilateral transfers between connected agents, incomes can be arbitrarily redistributed within any component of the network. As the main focus of our paper is network formation, to keep the model tractable we abstract away from enforcement constraints, and analogously to Bramoulle and Kranton (2007a, 2007b), we assume that all neighboring agents share risk efficiently, which in turn

---

\(^{15}\)This specification implies that we cannot impose a lower bound on the set of feasible consumption levels. As we show below, our framework readily generalizes to arbitrary income distributions, but the assumption of normally distributed shocks simplifies the analysis considerably.

\(^{16}\)It is well-known that for a vector of random variables not all combinations of correlations are possible. We implicitly assume that our parameters are such that the resulting correlation matrix is positive semidefinite.
leads to ex ante Pareto-efficient risk sharing at the level of each connected component.\footnote{For a model of informal insurance in social networks in which the set of feasible agreements are constrained by enforceability requirements, but in which the social network is exogenously given, see Ambrus et al. (2012).} While in practice risk sharing is imperfect, prefect risk sharing provides a useful benchmark. It is also straightforward to extend the model so that some income is publicly observed and perfectly shared while remaining income is privately observed and never shared. Results are very similar for this more general setting.\footnote{Kinnan (2011) finds evidence that hidden income can explain imperfect risk sharing in Thai villages relative to the enforceability and moral hazard problems we are abstracting from. Cole and Kocherlakota (2001) show that when individuals can privately store income, state-contingent transfers are not possible and risk sharing is limited to borrowing and lending.}

Formally, a risk-sharing agreement specifies a consumption vector $c$ for every state, in a way that $\sum_{i \in C} c_i(\omega) \leq \sum_{i \in C} y_i(\omega)$ for every state $\omega$ and network component $C$.

In Proposition 1 below we show that the CARA utilities framework has the convenient property that expected utilities are transferrable, in the sense defined by Bergstrom and Varian (1985). Moreover, ex ante Pareto efficiency is equivalent to minimizing the sum of variances, and it is achieved by agreements that at every state split the sum of the incomes at each network component equally among members and then adjust these shares by state-independent transfers. The latter determine the division of the surplus created by the risk-sharing agreement. We emphasize that this result does not require any assumption on the distribution of incomes, only that agents have CARA utilities.

**Proposition 1** For CARA utility functions certainty equivalent units of consumption are transferrable across agents, and if $L(S)$ is a network component the Pareto frontier of ex ante risk-sharing agreements among agents in $S$ is represented by a simplex in the space of certainty equivalent consumption. The ex ante Pareto-efficient risk-sharing agreements for agents in $S$ are those that satisfy:

$$\min \sum_{i \in S} \text{Var}(c_i) \quad \text{subject to} \quad \sum_{i \in S} c_i(\omega) = \sum_{i \in S} y_i(\omega) \quad \text{for every state} \; \omega,$$

and they are comprised of agreements of the form

$$c_i(\omega) = \frac{1}{|S|} \sum_{k \in S} y_k(\omega) + \tau_i \quad \text{for every} \; i \in S \; \text{and state} \; \omega.$$

**Proof.** See Appendix A. \hfill \blacksquare

The first statement in Proposition 1 can be established by showing that, with CARA utilities, the certainty equivalent consumption for a lottery is independent of the consumption level. The results on the Pareto-efficient risk-sharing arrangements can be obtained by applying the classic Borch rule (Borch (1962), Wilson (1968)) and algebraically manipulating the resulting conditions.

Proposition 1 implies that the total surplus generated by an efficient risk-sharing arrangement is an increasing function of the reduction in the sum of consumption variances. For general distribution of shocks, this function can be complicated. However, when shocks are jointly normally distributed, then $c_i = \frac{1}{|S|} \sum_{k \in S} y_k + \tau_i$ is also normally distributed, and
E(u(c_i)) = E(c_i) - \frac{1}{2} \text{Var}(c_i).^{19} \text{ Hence in this case the total social surplus generated by efficient risk-sharing agreements is proportional to the total consumption variance reduction. This greatly simplifies computing surpluses in the analysis below.}

We use \(TS(L)\) to denote the expected total surplus generated by an ex ante Pareto efficient risk-sharing agreement on network \(L\), relative to agents consuming in autarky.

**Division of Surplus**

We assume that agents on a connected component divide the total surplus created by the risk-sharing arrangement according to the Myerson value (Myerson (1977), (1980)). The Myerson value is a cooperative solution concept defined in transferable utility environments, that is a network-specific version of the Shapley value. The basic idea behind it is the same as for the Shapley value. For any order of arrivals of players, the incremental contribution of an agent to the total surplus can be derived as the difference between the total surpluses generated by the subgraph of \(L\) defined by the given agent and those who arrived earlier, and by the subgraph that is defined by only those agents who arrived earlier. It is easy to see that, for any order of arrivals, this way the total surplus generated by \(L\) gets exactly allocated to the set of all agents. The Myerson value then allocates the average incremental contribution of a player to the total surplus, taken over all possible orders (permutations) of players, as the player’s share of the total surplus. So agent \(i\)’s Myerson value is:

\[
\text{MV}_i(L) \equiv \frac{\sum_{S \subseteq N} (|S| - 1)! (n - |S|)!}{n!} \left( TS(L(S)) - TS(L(S/i)) \right).
\]

Our motivation for using the Myerson value is twofold. First, if agents on a connected component decide on the division of the surplus in a centralized manner, the Myerson value is selected based on normative considerations: the benefits received by an agent from the agreement should be equal to the average contribution of the agent to the social surplus. As shown in Myerson (1980), it is also implied by two very basic axioms: efficiency at the component level, and the requirement that the marginal benefit of a link is the same for the connecting agents, labeled as balanced contributions. Second, as we show below, a simple decentralized procedure involving bilateral transactions also selects the Myerson value.

To start with, consider a procedure proposed in Bramoullé and Kranton (2007a): after the realization of endowments, neighboring agents have repeated meetings with each other, in an arbitrary order, and each time they equalize their incomes. As shown in the above paper, such a procedure leads to splitting the total endowment at any component of the network equally among agents in the component. We increment this procedure with an ex ante stage, in which neighboring agents make bilateral agreements on ex ante transfers that are state independent and not subject to ex post redistribution. Analogously to Stole and Zwiebel (1996), we require these transfers to be renegotiation proof.

Formally, for network \(L\), let \(u_i(L)\) be \(i\)’s expected payoff ex ante (before incomes are realized). We assume that for every \(l_{ij} \in L\), agents \(i\) and \(j\) meet to negotiate a transfer

\(^{19}\text{See for example Arrow (1965).}\)

\(^{20}\text{Our assumption that there is perfect risk sharing among path connected agents ensures that a coalition of path connected agents generates the same surplus regardless of the exact network structure connecting them. This means that we are in the communication game world originally envisaged by Myerson. We do not require the generalization of the Myerson value to network games proposed in Jackson and Wolinsky (1996), which somewhat confusingly is also commonly referred to as the Myerson value. Jackson (2006) critiques the generalization of the Myerson value on the grounds that the allocation is insensitive to heterogeneities in people’s surplus generating capabilities that are captured by alternative, unformed networks. These heterogeneities are not possible in communication games.}\)
before the endowments realize. We let each agent have the option of holding up the other by deleting the link, and then negotiating its reformation. The network of social connections that have already formed, provides an upper bound on the set of relationships that can be used for the purposes of risk sharing, so new links cannot be formed, but existing links can be deleted and reformed. Implicitly, it takes a long time to establish a relationship, but once formed that relationship can be dissolved and possibly reformed much faster. We suppose that when a link is reformed the agents ‘split the difference’ and benefit equally from the link. Robustness to these renegotiations then requires that, at the margin, each formed link benefits both agents equally. Of course, in order to calculate what agents $i$ and $j$ would receive in the network without the link $l_{ij}$, we have to consider what would happen if links were renegotiated in the network without $l_{ij}$ and so on. The result is a recursive system of equations. The value of the link is only directly pinned down when it is the only link for both agents, and without the link both agents receive their autarky outcomes. Iterating, we can now consider networks with two links, and so on. At each stage of this recursion we require that for each link $l_{ij} \in L$, the incremental benefit provided by the link is split equally between agents $i$ and $j$. Following the terminology of Stole and Zwiebel (1996), we label risk-sharing arrangements satisfying the resulting criterion as robust to renegotiation.

Formally, for any network $L$, let $U(L)$ be the set of mappings from all subnetworks of $L$ to $\mathbb{R}^N$, representing payoff vectors to agents for different network realizations. We refer to elements of $U(L)$ as contingent payoff schemes given $L$. For a contingent payoff scheme $u$, let $u_i(L')$ denote the payoff of agent $i$ given $L' \subseteq L$. Lastly, for ease of exposition, we use $L - l_{ij}$ instead of $L \{l_{ij}\}$ for the network obtained from $L$ by deleting link $l_{ij}$.

**Definition 2** For any network $L$, the payoff vector $(u_1, ..., u_N)$ is robust to renegotiation if there is a contingent payoff scheme $u$ given $L$ such that:

(i) $u_i(L') - u_i(L' - l_{ij}) = u_j(L') - u_j(L' - l_{ij})$ for every $i, j \in N$ and $L' \subseteq L$;

(ii) $\sum_{i \in N} u_i(L') = TS(L')$ for every $L' \subseteq L$.

Below we show that the requirement of robustness to split-the-difference renegotiation implies all agents must receive their Myerson values.

**Proposition 3** For any network $L$, there is a unique vector of payoffs that is robust to renegotiation and at this outcome all agents receive their Myerson Values: $u_i(L) = MV_i(L)$.

**Proof.** We will use the axiomatization of the Myerson value established in Myerson (1980).\(^{21}\) This axiomatization states that there exists a unique payoff division rule satisfying that (i) any link benefits the two connecting agents equally; and (ii) the outcome is efficient at the component level. Formally, these requirements are equivalent to properties (i) and (ii) in definition 2. So a payoff vector is robust to renegotiation if and only if it corresponds to the Myerson value. Since the Myerson value is unique, there is a also a unique payoff vector that is robust to renegotiation.

\(^{21}\)An assumption made by Myerson when defining the Myerson value is that the surplus generated by a connected component is independent of the network structure within the component. Networks are viewed as communication structures. Although the Myerson value was redefined more generally in Jackson and Wolinsky (1996), as we apply Myerson’s axiomatization it is important that our setting satisfies the communication structure assumption.
Proposition 2 is a direct implication of Myerson’s axiomatization of the value. A special case of Proposition 2 is Theorem 1 of Stole and Zwiebel (1996), which restrict attention to a star network. Our contribution is to point out that the connection between robustness to renegotiation and the Myerson value applies across all networks.

3 Investing in Social Relationships

In this section we first formally introduce a game of network formation in which establishing links is costly, and define the concepts of efficient networks and different types of inefficiencies in network formation.

We consider a two-period model in which in period 1 all agents simultaneously choose which other agents they would like to form links with, and in period 2 agents agree upon the ex ante Pareto efficient risk-sharing agreement specified in the previous section (i.e., the total surplus from risk-sharing is distributed according to the Myerson value), whichever network forms in the first period.

Formally, in period 1, we consider a network formation game along the lines of Myerson (1991): all agents simultaneously choose a subset of the other agents, indicating who they would like to form links (relationships) with. A link is formed between two agents if and only if they both want to form it – i.e. if both agents select each other. When agent $i$ forms a link, he pays a cost $\kappa_w > 0$ if the link is with someone in the same group and $\kappa_a > \kappa_w$ if the link is with someone from a different group. This specification assumes that two agents forming a link have to pay the same cost for establishing the link. However, all of our results below would remain valid if we allowed the agents to share the total costs of establishing a link arbitrarily (namely if we allowed the agents in the first period not only indicating who they would like to establish links with, but also propose a division of the costs of establishing each link; a link would then only form if both agents indicate each other and they propose the same split of the cost). This is because for any link, the Myerson value rewards the two agents establishing the link symmetrically. Hence the agents can find a split of the link formation cost such that establishing the link is profitable for both of them if only if it is profitable for both of them to form the link with an equal split of the cost. For this reason we stick with the simpler model with exogenously given costs.

The collection of links formed in period 1 becomes social network $L$.

Normalizing the utility from autarchy to 0, agent $i$’s net payoff if network $L$ forms is:

$$U_i(L) = MV_i(L) - |N_{G(i)}(i; L)|\kappa_w - \left(|N(i; L)| - |N_{G(i)}(i; L)|\right)\kappa_a.$$

The solution concept we apply to the simultaneous move game described above is pairwise stability. A network $L$ is pairwise stable with respect to payoff functions $\{u_i(L)\}_{i \in N}$ if and only if for all $i, j \in N$, (i) if $l_{ij} \in L$ then $u_i(L) - u_i(L/\{l_{ij}\}) \geq 0$ and $u_j(L) - u_j(L/\{l_{ij}\}) \geq 0$; and (ii) if $l_{ij} \notin L$ then $u_i(L \cup l_{ij}) - u_i(L) > 0$ implies $u_j(L \cup l_{ij}) - u_j(L) < 0$. In words, pairwise stability requires that no two players can both strictly benefit by establishing an extra link with each other, and no player can benefit by unilaterally deleting one of his links.

From now on we refer to pairwise stable networks simply as equilibrium networks. Existence of a pairwise stable networks in our model follows from a result in Jackson (2003),

\[\text{\footnotesize{\cite{pin}}}
\]

For a complementary treatment of network formation when surplus is split according to the Myerson Value, see Pin (2011).
stating that whenever payoffs in a simultaneous-move network formation game are determined based on the Myerson value, there exists a pairwise stable network.

**Proposition 4 (Jackson, 2003)** There exists an equilibrium in the game of network formation.

A network \( L \) is efficient when there is no other network \( L' \) and no risk sharing agreement on \( L' \) that can make everyone at least as well off as they were on \( L \) and someone strictly better off. Let \( |L_w| \) be the number of within group links and let \( |L_a| \) be the number of across group links. As expected utility is transferable in certainty equivalent units, efficient networks must maximize the net total surplus \( NTS(L) \):

\[
NTS(L) \equiv CE\left(\Delta \text{Var}(L, \emptyset)\right) - 2|L_w|\kappa_w - 2|L_a|\kappa_a,
\]

where, for \( L' \subset L \), \( \Delta \text{Var}(L, L') \) is the additional variance reduction obtained by efficient risk sharing on network \( L \) instead of \( L' \), and \( CE(\cdot) \) denotes the certainty equivalent value of a variance reduction.

Clearly two necessary conditions for a network to be efficient are that the removal of a set of links does not increase \( NTS(L) \) and the addition of a set of links does not increase \( NTS(L) \). If there exists a set of links, the removal of which increases \( NTS(L) \) we will say there is overinvestment inefficiency. If there exists a set of links the addition of which increases \( NTS(L) \) we will say there is underinvestment inefficiency.\(^{23}\)

We will say that a link \( l_{ij} \) is essential if after its removal \( i \) and \( j \) are no longer path connected.

**Remark 5** Preventing overinvestment requires that all links are essential. Additional links create no social surplus and are costly. In all efficient networks, every component must therefore be a tree.

In most of the analysis below we focus on investigating the relationship between equilibrium networks and efficient networks.

### 4 Local Network Formation

In this section we assume that \( m = 1 \), that is agents are ex ante symmetric, and any differences in their outcomes stem from their equilibrium positions on the social network. This will lay the foundations for the more general case considered in the next section.

The social value of a non-essential link is 0. We begin the analysis by characterizing the social value of an essential link. As shown in the previous subsection, the social value of a link is proportional to the reduction it implies, through a Pareto efficient risk-sharing agreement, in the sum of the consumption variances. Let \( L(S_1) \) and \( L(S_2) \) be the network components of agent \( i \) and agent \( j \) on network \( L/\{l_{ij}\} \), and let \( |S_1| = s_1 \) and \( |S_2| = s_2 \). Then the sum

\(^{23}\)Note that these definitions are not mutually exclusive (there can be both underinvestment and overinvestment inefficiency) or collectively exhaustive (inefficient networks can have neither underinvestment nor overinvestment inefficiency if an increase in net total surplus is possible by the simultaneous addition and removal of edges).
of consumption variances on $L(S_1)$ and $L(S_2)$, assuming Pareto-efficient risk sharing, are $\frac{s_1 + s_1(s_1 - 1)\rho_w}{s_1} \sigma^2$ and $\frac{s_2 + s_2(s_2 - 1)\rho_w}{s_2} \sigma^2$, respectively. Once $S_1$ and $S_2$ are connected through $l_{ij}$, the sum of consumption variances on $L(S_1 \cup S_2)$ becomes $\frac{s_1 + s_2 + (s_1 + s_2)(s_1 + s_2 - 1)\rho_w}{s_1 + s_2} \sigma^2$. This implies that the consumption variance reduction induced by the link $l_{ij}$ is $\Delta \text{Var}(L \cup \{l_{ij}\}, L) = (1 - \rho_w)\sigma^2$. This means that the variance reduction, and therefore the surplus created by an essential link, in case of homogenous agents, is independent of the sizes of the components the essential link connects. Intuitively, an increase in the size of one of the components, say $s_1$, has two effects. On the one hand it increases the consumption variance reduction for agents $S_2$ when they get linked to agents $S_1$, as agents in the latter component can spread the income risk from agents $S_2$ more effectively. On the other hand it decreases the consumption variance reduction of agents $S_1$ when they get linked to agents $S_2$, as the increase in $s_1$ implies that risk-sharing is better on $S_1$ already. What the above formula shows is that for homogenous agents these two effects perfectly cancel each other out.

This implies that it is particularly simple to determine the gross surplus created by network $L$. Let $f(L)$ be the number of network components on $L$. Then:

$$CE\left(\Delta \text{Var}(L, \emptyset)\right) = (N - f(L))\frac{\lambda}{2}(1 - \rho_w)\sigma^2.$$ 

Since the surplus created by any essential link is $V \equiv \frac{\lambda}{2}(1 - \rho_w)\sigma^2$, total gross surplus is equal to the latter constant times the number of network component reductions relative to the empty network.

Next we investigate private incentives for link formation. Recall that the share of the surplus created by risk-sharing allocated to an agent $i$ is equal to the average incremental surplus created by adding him to the network, over all possible orders of arrival of players. Every link an agent has result in component reduction of 1 or no component reduction when $i$ is added. The link $l_{ij}$ will reduce the number of components in the graph by one when $i$ is added, and only if $j$ has already been added and there is no other path between $i$ and $j$. If there is another path between $i$ and $j$, or $j$ has not been added yet, then the link $l_{ij}$ will not result in any component reduction when $i$ is added. Suppose agents $S$ have been added before $i$ for a given permutation of arrival orders. As before we let $L(S) \subseteq L$ be the subgraph of $L$ such that $l_{ij} \in L(S)$ if and only if $l_{ij} \in L$, $i \in S$ and $j \in S$. If $j \not\in S$, then the link $l_{ij}$ is not formed when $i$ is added and so $i$ receives no benefit from it. If $j \in S$ then $l_{ij}$ reduces the number of components present in the graph by 1 if there is no other path from $i$ to $j$ on $L(S \cup i)$ and reduces the number of components in the graph by zero otherwise. The link $l_{ij}$ is valuable to $i$ when added if and only if it is essential on the graph $L(S \cup i)$.

**Remark 6** Let $L^{e}(S \cup i) \subseteq L(S \cup i)$ be the set of essential links on $L(S \cup i)$. The incremental surplus generated when $i$ is added is proportional to the number of essential links $i$ has on $L(S \cup i)$. In other words:

$$CE\left(\Delta \text{Var}(L(S \cup i), L(S))\right) = \left|L(i) \cap L^{e}(S \cup i)\right| \cdot V.$$ 

We now characterize the set of pairwise stable networks. Some additional terminology will be helpful. A minimal path between $i$ and $j$ is any path between $i$ and $j$ such that no other path between $i$ and $j$ is a subsequence. If there are $K$ minimal paths between $i$ and $j$ on the network $L$, we let $P(i, j, L) = \{P_1(i, j, L), \ldots, P_K(i, j, L)\}$ be the set of these paths.
We let $|P_k(i, j, L)|$ be the cardinality of the set of different agents in the sequence $P_k(i, j, L).$  
We can now use these definitions to define a quantity that captures how far away two agents are on a network in terms of the probability that they will be connected without a direct link when the second of the two agents is added in a random permutation. We will refer to this distance as the agents’ Myerson distance:

$$md(i, j, L) = \frac{1}{2} - \sum_{k=1}^{\frac{|P(i,j,L)|}{2}} (-1)^{k+1} \sum_{1 \leq i_1 < \cdots < i_k \leq |P(i,j,L)|} \left( \frac{1}{|P_{i_1} \cup \cdots \cup P_{i_k}|} \right).$$

This expression calculates the probability that for a random permutation the link $l_{ij}$ will be essential immediately after $i$ is added, using the classic inclusion-exclusion principle from combinatorics. This probability is important because it affects $i$’s incentives to link to $j$.

As an illustration, suppose that there is a unique indirect path $P_1(i, j, L)$ between $i$ and $j$ that contains $K$ agents, including $i$ and $j$. We then have $md(i, j, L) = 1/2 - 1/K$. To see where this expression comes from, note that there are two reasons why $l_{ij}$ might not be essential when $i$ is added. First, $j$ might not yet have been added. This occurs with probability $1/2$. A second way that $l_{ij}$ might not be essential is if all other agents on the path $P_1(i, j, L)$, including $j$, are added before $i$. This occurs with probability $1/K$. Combining these probabilities is straightforward. The probability of both events occurring is 0 because collectively they require $j$ to be both present and absent; so, we can just sum them. Thus, the probability that $l_{ij}$ is essential when $i$ is added, is $1 - 1/2 - 1/K = md(i, j, L)$.

Suppose now that there are two paths between $i$ and $j$, $P_1(i, j, L)$ and $P_2(i, j, L)$, on the network $L$. Suppose that $P_1(i, j, L) = \{i, i’, i'', j\}$ and $P_2(i, j, L) = \{i, i’, i’’, j\}$. We need to find the probability that either of these paths are present when $j$ is added. To avoid double counting, we need to add the probability that $P_1(i, j, L)$ is present (1/4) to the probability $P_2(i, j, L)$ is present (1/4) and then subtract the probability that both are present (1/5). So $md(i, j, L) = 1 - 1/2 - 1/4 - 1/4 + 1/5$. The Myerson distance calculation provides the general way of accounting for the probability that at least one of multiple possible paths is present.

**Proposition 7** If agents are ex-ante homogeneous $(m = 1)$, a network $L$ is pairwise stable if and only if

(i) $md(i, j, L \setminus \{l_{ij}\}) \geq \kappa_w/V$ for all $l_{ij} \in L$, and

(ii) $md(i, j, L) \leq \kappa_w/V$ for all $l_{ij} \notin L$.

**Proof.** See Appendix A. □

The first step in the proof of Proposition 7 establishes that the value of a link $l_{ij}$ to $i$ (and $j$) is $V$ if the link is essential when $i$ is added and 0 otherwise. Suppose a link $l_{ij}$ is essential on $L$. It will then always induce a component reduction of one, for any orders of arrival, when the later of $i$ and $j$ are added. So $md(i, j, L) = 1/2$ and $l_{ij}$ will be formed as long as $V > 2\kappa_w$. As $V$ is the social value of forming the link and $2\kappa_w$ is the total cost of forming it, with homogeneous agents there is never underinvestment in equilibrium. This argument is formalized in Proposition 8.
Proposition 8 If all agents are homogenous then there is never underinvestment in equilibrium. Furthermore, there is never overinvestment in an essential link.

Proof. For there to be underinvestment in a pairwise stable network \( L \), there must exist a link \( l_{ij} \notin L \) for which the social value is created than the cost of formation, so that \( TS(L \cup l_{ij}) - TS(L) > 2 \kappa_w \).

As non-essential links have no social value, \( l_{ij} \) must be essential on \( L \cup l_{ij} \) and so \( TS(L \cup l_{ij}) - TS(L) = V \) and \( md(i, j, L) = 1/2 \). By Proposition 7, as \( l_{ij} \) is not formed and the network is pairwise stable, \( md(i, j, L) \leq \kappa_w / V \) and so \( TS(L \cup l_{ij}) - TS(L) \leq 2 \kappa_w \), which is a contradiction.

Combining the results above reveals the following properties of equilibrium with homogeneous networks.

Corollary 9 For homogeneous agents, if \( 2 \kappa_w > V \) then the only stable network is the empty one and this network is efficient, while if \( 2 \kappa_w < V \) then all equilibrium networks have only one network component (all agents are path-connected).

Corollary 10 For homogeneous agents, in any efficient equilibrium \( u_i = |N(i; L)|(V/2 - \kappa_w) \) and agents’ payoffs are proportional to their degree centralities.

Motivated by Corollary 9, and our data in which the observed networks are clearly not empty, we will assume from now on that that \( 2 \kappa_w < V \). We will refer to this as our regularity condition.

Although, as shown above, with homogenous agents there is never overinvestment in essential links, there can be overinvestment in the form of superfluous links. Moreover, if the cost of establishing a link is low enough, such inefficiencies are unavoidable. The reason is that although a superfluous link \( l_{ij} \) does not create any social surplus, it always increases the Myerson value of the participants, through increasing their incremental contributions for some orders of arrivals (for example when \( i \) and \( j \) are the first two arrivers).

Since for homogeneous agents underinvestment is never an issue, but overinvestment can be, in what follows we focus on investigating what network structures minimize incentives for overinvestment. As we will see, this question is also related to the issue of inequality that different network structures imply. For concreteness, we define inequality on network \( L \) to be the range in payoffs, that is the difference between the maximum and minimum expected payoff implied by \( L \).

On any tree network with three or more nodes, there must exist leaf nodes that have degree 1 and non-leaf nodes that have degree 2 or higher. By Corollary 10, a lower bound on inequality is therefore \((V - 2 \kappa_w)/2\). Moreover, for any tree network with \( n \) nodes there are exactly \( n - 1 \) links and so all agents must have degree \( n - 1 \) or lower. This means that an upper bound on inequality is \((n - 2)(V - 2 \kappa_w)/2\).

Let the line network be the unique (tree) network, up to a relabeling of agents, in which there is a path from one (end) agent to the other (end) agent that passes through all other agents exactly once (see Figure 1a). Let the star network be the unique tree network, up to a relabeling of agents, in which one (center) agent is connected to all other agents (see
Figure 1: Stable and efficient networks for $2\kappa_w \in \left(\frac{2}{3}V, V\right)$, where $V$ is the social value of an essential link.

Figure 1c). On all tree networks connecting at least three agents there are some agents who have degree 1 (leaf nodes) and some agents that have degree greater than 1 (branch nodes). As the line network achieves the lower bound on inequality it is the most equitable efficient network, while the star achieves the upper bound on inequality and so is the least equitable efficient network.

Recall that on any efficient network, there is a unique path between any two connected agents. The private incentives of two agents to form a superfluous link only depends on the length of the path connecting them, in a strictly increasing manner. Suppose $d$ is the number of agents between $i$ and $j$ in the unique path connecting them. The probability that this path exists when agent $i$ is added to the network in the Myerson distance calculation is $1/d$. In addition, if agent $j$ has not yet been added, which occurs with probability $1/2$, $i$ would not benefit from the link $l_{ij}$. So, $i$’s expected payoff from forming a superfluous link to $j$ is $(1 - 1/2 - 1/d) V$. As $d$ gets large this converges to $V/2$ which is the value $i$ receives from forming an essential link. These claims are formalized in the next proposition.

Recall that $d(L)$ is the diameter of a network $L$.

**Proposition 11** For homogeneous agents, if $L$ is efficient then:

(i) As $d(L)$ gets large, there exists a superfluous potential link for which the incentives to add this link converges to the incentives to add an essential link.

(ii) $L$ is stable if and only if its diameter is less than $\overline{d}(2\kappa_w)$, where $\overline{d}(\cdot)$ is increasing and integer-valued.

**Proof.** Recall that for a network $L$ with diameter $d(L)$, there exists agents $i$ and $j$ for whom the length of the unique path connecting them is $d(L)$. Consider the incentives of these agents to form the link $l_{ij}$. By Lemma 7, $i$ and $j$ will want to form the link if and only if:

$$\frac{V - 2\kappa_w}{V} \geq 2 \left( \frac{1}{d(L)} \right).$$

As $d(L)$ gets large the left hand side converges to 0 and so in the limit, the condition for a link to be formed becomes $V \geq 2\kappa_w$, which is the condition for an essential link to be formed.

By Proposition 8 there is never any underinvestment in an efficient networks $L$. An efficient network will then be stable if and only if there are no incentives to form a superfluous
link. As two agents’ incentives to form a superfluous link are increasing in the path length between them, \( L \) is stable if and only if:

\[
d(L) \leq 2 \left( \frac{V}{V - 2\kappa_w} \right).
\]

Setting \( d(2\kappa_w) = \lceil 2V/(V - 2\kappa_w) \rceil \) complete the proof. \( \square \)

The following corollary of the previous result reveals a novel trade-off between maximizing efficiency and decreasing inequality.

**Corollary 12** For homogeneous agents, if there exists an efficient equilibrium network then star networks are equilibrium networks. Moreover, for a range of cost parameters for establishing a link within group, the only efficient equilibrium networks are stars.

Hence the star, which is the efficient network that maximizes inequality, is also the most stable, as it minimizes agents’ incentives to establish superfluous links. Conversely, the line, which minimizes inequality in the class of efficient networks, also maximizes the diameter of the network and so is the efficient network that is most unstable (it is stable for the smallest set of linking cost parameters among all efficient networks).

### 5 Connections Across Groups

We now generalize our model by permitting multiple groups. These different groups might correspond to people from different villages, different occupations or from different social status groups, such as castes. We will first show that under our regularity condition there is still never any underinvestment within group. However, this does not apply to links that bridge groups. As, by assumption, incomes are more correlated within group than across group, there can be significant benefits from establishing such links and not all these benefits accrue to the agents forming the link. Intuitively, an agent establishing a bridging link to another group provides other members of her group with access to a less correlated income stream, which benefits them. As agents providing such bridging links are unable to appropriate the benefits these links generate, and these links are relatively costly to establish, there can be underinvestment.

To analyze the incentives to form links within group, we first need to consider the variance reduction obtained by a within group link. Such a link may now connect two otherwise separate components comprised of arbitrary distributions of agents from different groups. Suppose that agents \( S_0 \cup \cdots \cup S_k \) and agents \( \widehat{S}_0 \cup \cdots \cup \widehat{S}_k \) form two distinct network components, where for every \( i \in \{0, \ldots, k\} \), agents in \( S_i \) and in \( \widehat{S}_i \) are all from group \( i \) \( (i = 0, \ldots, k) \). Consider now a potential link \( l_{ij} \) connecting the two otherwise disconnected components. The variance reduction obtained is:

\[
\Delta \Var(L \cup l_{ij}, L) = \Var(L(S_0, \ldots, S_k)) + \Var(L(\widehat{S}_0, \ldots, \widehat{S}_k)) - \Var(L(S_0 \cup \widehat{S}_0, \ldots, S \cup \widehat{S}_k)).
\]

Recalling that,
\[
\text{Var}(L(S_0, S_1, \ldots, S_k)) = \frac{k \sum_{i=0}^{k} (s_i + s_i(s_i - 1)\rho_w) + 2\rho_w \sum_{i=0}^{k-1} \sum_{j=i+1}^{k} s_j}{\sum_{i=0}^{k} s_i} \sigma^2,
\]

some algebra yields\(^{26}\)

\[
\Delta \text{Var}(L \cup l_{ij}, L) = \left[ (1 - \rho_w) + \frac{\sum_{i=0}^{k} \left( \hat{s}_i \sum_{j=0}^{k} s_j - s_i \sum_{j=0}^{k} \hat{s}_j \right)^2}{\left( \sum_{i=0}^{k} s_i \right) \left( \sum_{i=0}^{k} \hat{s}_i \right) \left( \sum_{i=0}^{k} s_i + \hat{s}_i \right)} (\rho_w - \rho_a) \right] \sigma^2. \tag{2}
\]

The key feature of this equation is that it is always weakly greater than \((1 - \rho_w)\sigma^2\), which is the variance reduction we found in the previous section when all agents were from the same group. So, the presence of across group links within a group only increases the incentives for within group links to be formed. This implies that there will still be no under-investment as long as our regularity condition is met and \(2\kappa_w \leq V = \frac{1}{2} (1 - \rho_w) \sigma^2\). Recall that this regularity condition just requires that it is efficient for two agents, both without any other connections and in the same group to engage in risk-sharing.

**Proposition 13** Under the regularity condition that \(2\kappa_w \leq V\), there will no under-investment between any two agents from the same group in any stable network.

Motivated by Proposition 13, within group we will continue to focus on the problem of overinvestment rather than underinvestment. However, in contrast to Proposition 13, there can be underinvestment across group. The key insight is that, as opposed to the case of homogenous agents, where the value of an essential link does not depend on the sizes of the components it connects, the value of an essential link connecting two different groups of agents increases in the sizes of the components. To show this formally, consider an isolated group that has no across group connections and consider the incentives for a first such connection to be formed. Let the first component consists of agents from a single group, say group 0, and the second component consists of agents from any other groups (1 to \(k\)). The variance reduction obtained by connected these two components then simplifies to:

\[
\Delta \text{Var}(L \cup l_{ij}, L) = \left[ (1 - \rho_w) + \frac{\hat{s}_0 \left( \sum_{i=1}^{k} s_i \right)^2 + \sum_{i=1}^{k} s_i^2}{\sum_{i=1}^{k} s_i \left( \hat{s}_0 + \sum_{i=1}^{k} s_i \right)} (\rho_w - \rho_a) \right] \sigma^2, \tag{3}
\]

\[
\frac{\partial \Delta \text{Var}(L \cup l_{ij}, L)}{\partial \hat{s}_0} = \frac{\left( \sum_{i=1}^{k} s_i \right)^2 + \sum_{i=1}^{k} s_i^2}{\left( \hat{s}_0 + \sum_{i=1}^{k} s_i \right)^2} (\rho_w - \rho_a) \sigma^2 > 0. \tag{4}
\]

The inequality follows since \(\rho_w > \rho_a\).

An immediate implication is that agents \(i\) and \(j\), who connect two otherwise unconnected groups receive a strictly smaller combined private benefit than the social value of the link.

\(^{26}\)One of the key steps to simplifying the expression is noting that: \(2 \sum_{i=0}^{k-1} \left( s_i \sum_{j=i+1}^{k} s_j \right) = \left( \sum_{i=0}^{k} s_i \right)^2 - \sum_{i=0}^{k} s_i^2\).
To see why, consider the Myerson Value calculation. In most orders of arrivals when the second agent of the pair \( ij \) arrives, not all other agents on the components of \( i \) and \( j \) have arrived yet. Hence, for most orders of arrivals the incremental contribution of the link to the Myerson values of the connecting agents is smaller than the social value of the link. For the remaining orders of arrival the incremental contribution of the link the Myerson value is its social value. Averaging over these orders of arrivals, the link contributes less to the Myerson values of \( i \) and \( j \) than its social value leading to the possibility of underinvestment.

We formalize the resulting possibility of underinvestment in Proposition 14.

Let \( S_g = \{ i : G(i) = g \} \) denote the agents in group \( g \).

**Proposition 14** If \( m \geq 2 \) then underinvestment is possible in equilibrium.

**Proof.** We will show that if there are \( m \geq 2 \) equal sized groups then there is a range of parameters \( \kappa_w > 0 \) and \( \kappa_a > \kappa_w \) such that in any equilibrium all groups are disconnected, despite an extra link connecting any two groups having a strictly positive social value. By assumption the within group correlation coefficient \( \rho_w < 1 \) and the coefficient of absolute risk aversion \( \lambda > 0 \). Together, these parameter restrictions imply that the certainty equivalent value of a variance reduction from connecting one agent to any group of other agents is strictly positive. So, for \( \kappa_w \) sufficiently close to 0, in all equilibria any two agents from the same group must be path connected.

Assume now that all groups form separate network components, and consider a potential extra link \( l_{ij} \) connecting groups \( g \) and \( g' \). As shown above, the change in total variance, and so surplus, achieved by connecting agents in \( S_g \) to agents in \( S_{g'} \) is increasing in the size of both groups, \( s_g \) and \( s_{g'} \) respectively. This means that the marginal contribution to total surplus of the link \( l_{ij} \) is higher than the contribution to total surplus when the later of \( i \) and \( j \) are added to the network, unless \( i \) or \( j \) is added last. This implies that \( MV(i; L \cup l_{ij}) - MV(i; L) < TS(L \cup l_{ij}) - TS(L) \) for all \( i \in S_g \) and \( MV(j; L \cup l_{ij}) - MV(j; L) < TS(L \cup l_{ij}) - TS(L) \) for all \( j \in S_{g'} \). There thus exists a range of \( \kappa_a \) for which \( MV(i; L \cup l_{ij}) - MV(i; L) < \kappa_a \) and \( MV(j; L \cup l_{ij}) - MV(j; L) < \kappa_a \), but \( TS(L \cup l_{ij}) - TS(L) > 2 \kappa_a \). For such parameters, there is an equilibrium in which within groups agents are completely connected, but there is no link across groups, despite it being socially desirable. ■

Besides underinvestment, overinvestment is also possible across group. Forming superfluous links will increase an agent’s share of surplus without improving overall risk sharing and can therefore create incentives to overinvest. Nevertheless, when \( \kappa_a \) is relatively high, underinvestment rather than overinvestment in across group links will be the main efficiency concern. In many settings within group links are relatively cheap to establish in comparison to across group links. For example, when the different groups correspond to different castes, as in our data, it can be quite costly to be seen to interact with members of the other caste (e.g., Srinivas (1962), Banerjee et al. (2013b)). Motivated by this, and because across group links are considerably sparser in our data (to be described in the next section) than within group links, we focus our attention on this parameter region. More concretely, below we investigate what within-group network structures create the best incentives to form across group links and what network structures minimize the incentives for overinvestment within group. Remarkably, we will find that these two forces push local network structures in the same direction, and in both cases towards inequality in the society.

We begin by considering local overinvestment within groups, which corresponds to the forming of superfluous within group links. We found in the previous section that, for homoge-
ous agents, the star was the efficient network that minimized the incentives for overinvestment. However, once we include links to other groups the analysis is more complicated. The variance reduction a within group link generates is still zero if the link is superfluous, but when the link is essential it now depends on the distribution of agents across the different groups the link grants access to. Moreover, the variance reduction may be decreasing or increasing in the number of people in a specific such group.\textsuperscript{27} This makes Myerson value calculation substantially more complicated. With homogenous agents, all that mattered was whether the link was essential when added. Now, for each arrival order in which the link is essential, we also need to keep track of the distribution of agents across the different groups who are being connected. Nevertheless, our earlier result generalizes to this setting, although the argument establishing the result is more subtle.

**Proposition 15** The local network structure that minimizes the incentives to overinvest within group is the local star, with the center agent holding all across group links. If any other local network is robust to within group overinvestment, this network is also robust to within group overinvestment.

**Proof.** Equation 2 shows that a lower bound on the variance reduction obtained by an essential link connecting any two components is \( (1 - \rho_w)\sigma^2 \). This is the variance reduction obtained by an essential within group link on an isolated component. So, by the Myerson calculation, starting from an isolated component, adding any set of across group links weakly increases the incentives for agents to form superfluous within group links. In other words, we have found a lower bound on the incentives to overinvest within group.

We now show that the local star, with all across group links held by the center node, achieves this lower bound. The key insight is that the presence of the across group link does not increase the incentives for overinvestment within group. Consider two periphery nodes \( i, j \) in the same local star, and consider their incentives to form the superfluous link \( l_{ij} \). The Myerson value calculation implies that the agents forming this link receive the link’s average marginal contribution to total surplus across all permutations in which the agents can be added. A necessary condition for the additional link \( l_{ij} \) to be essential when \( i \) is added is that the central node has not yet been added. Otherwise, there is already a path from \( i \) to \( j \) (or \( j \) has not yet been added). Thus, for every possible permutation, the additional link \( l_{ij} \) increases \( i \)'s average marginal contribution to total surplus by exactly the same amount, regardless of whether the central agent has an across group link or not.

As by Corollary 12 the local star minimized overinvestment incentives absent the across-group link, and as incentives can only be increased by the addition of across group links, the local star (with all across group link held by the center agent) must minimize overinvestment incentives in the presence of across-group links. In other words, once the across-group links are added (to the center node), the incentives to overinvest within group are no higher for this network, but are weakly higher for all other efficient networks. \( \blacksquare \)

We now consider the local network structures that maximize the incentives for an across group link to be established. We have already established that the marginal contribution to total surplus of a first bridging link is increasing in the size of the groups it connects. It follows that the agents with the strongest incentives to form such links, are those who will

\textsuperscript{27}In the case of an essential across-group link that bridges two otherwise disconnected groups, the comparative statics are unambiguous. In this case, the variance reduction is increasing in the size of the groups connected, as shown by inequality (4).
be linked to most other agents within their group when they added to the network in the Myerson calculation. The result below formalizes this intuition.

Let $P(S_k)$ be the set of possible permutations for agents $S_k$. For any permutation $P \in P(S)$, let $T_i(P)$ be the set of agents $i$ is path connected to on $L(S')$ where $S'$ is the set of agents including $i$ drawn weakly before $i$. Let $T_i^{(m)}$ be a random variable equal to the cardinality of $T_i(P)$, conditional on $i$ being the $m$-th agent to be drawn according to $P$.

We will say that agent $i \in S_k$ is more central within group than agent $j \in S_k$, if $T_i^{(m)}$ first-order stochastically dominates $T_j^{(m)}$ for all $m \in \{1, 2, ..., |S_k|\}$. In other words, considering all the permutations in which $i$ is the $m$-th agent, and all the permutations in which $j$ is the $m$-th agent, the size of $i$'s component at $i$'s arrival is larger than that of $j$'s at $j$'s arrival in the sense of first-order stochastic dominance. This measure of centrality provides a partial ordering of agents.

**Proposition 16** Suppose agents in $S_0$ form a network component, and all other agents in $N$ form another network component. Let $i, i' \in S_0$ and let $j \not\in S_0$. If $i$ is more central within group than $i'$, then $i$ receives a higher payoff from forming $l_{ij}$ than $i'$ receives from forming $l_{i'j}$:

$$MV(i; L \cup l_{ij}) - MV(i; L) > MV(i'; L \cup l_{i'j}) - MV(i'; L)$$

**Proof.** See Appendix A. ■

The proof of Proposition 16 pairs arrival orders for a more central and less central agents so that in each case the more central agent, when added, is connected to weakly more people in the same group and the same set of people from other groups as the less central agent. Such a pairing of arrival orders is possible from the definition of centrality, and in particular the first order stochastic dominance it requires.

Proposition 16 shows that more central agents have better incentives to form intergroup links. We can then consider the problem of maximizing the incentives to form intergroup links by choosing the local network structures (networks containing only within group links). We will say that the local network structures that achieve these maximum possible incentives are most robust to underinvestment inefficiency across group.

**Corollary 17** The efficient local network structure most robust to underinvestment inefficiency across group is the star, with the potential across-group link holder at the center. If any other local network is robust to underinvestment across group, then so is the star.

**Proof.** Within a local network, an agent connected to all other agents is weakly more central than any other agent. For any permutation, such an agent is always connected to all other agents that are before him in the permutation. As no agent can ever be connected to an agent after them in a permutation upon being added to the graph, an agent connected to everyone else is more central than all other agents and so by Proposition 16 has stronger incentives to form an across group link.

Efficient networks are always trees and the star network is the unique tree network in which one agent is directly connected to all others. ■

---

28An alternative and equivalent definition is that $i$ is more central than $j$ if there exists a bijection $B : P(S_k) \to P(S_k)$ such that $|T_i(P)| \geq |T_j(B(P))|$ and $P(i) = P'(j)$, where $P(i)$ is $i$'s position in the permutation $P$ and $P' = B(P)$.
The above results establish that for two groups, the efficient networks that minimize within-group overinvestment and across-group underinvestment are center-connected stars, as in Figure 2.

Further, the above results further reinforce the tension between efficiency and equality. The local star not only minimizes the incentives for within-group overinvestment, it also minimizes the incentives for across-group underinvestment. If an agent $i$ provides an across-group link, then of all the possible local network structures, the star, with $i$ at the center, maximizes $i$’s payoff. Nevertheless, for some (but not all) parameter values, a local star will be more equitable when the central agent provides an across group link than without it, because the across-group link generates positive spillovers to the whole group.

6 Empirical Analysis

We now test the predictions of our model in the data. An extensive literature documents that in India, caste plays a significant role in informal risk-sharing. For instance, Morduch (1991, 1999, 2004) and Walker and Ryan (1990) show that informal insurance functions rather well within caste (though it still is imperfect) while there is very limited insurance across caste. Morduch (2004) discusses this literature at large, which primarily uses the methodology developed in Townsend (1994) to describe the extent of risk sharing within and across caste groups in a village. In the language of our model, every caste corresponds to a different group and it is ex ante more costly for agents to form cross-caste links. This is in line with extensive sociological literature (see e.g., Srinivas (1962)).

For our analysis, we make use of a unique and detailed social network dataset from 75 villages in Karnataka, India, particularly well-suited for our analysis as (i) it involves numerous independent villages (essential for inference, though most network-studies have a just one or a handful of villages), (ii) it includes complete network data across both financial and social connections across almost all households in every village (network-based studies are notoriously subject to measurement error), and (iii) caste is salient in these communities.

29By Corollary 15, the local star maximizes $i$’s incremental payoff when $i$ provides an across-group bridging link. The local star also maximizes the value of each within-group link $i$ has as all such links are essential for all arrival permutations. Moreover, each within-group link is valuable as under our regularity condition $V > 2\kappa_w$, and equation 2 shows that the value of an essential within-group link is at least $V$. 

22
6.1 Setting and Data

The data we use were collected by Banerjee, Chandrasekhar, Duflo and Jackson (2013, 2014) from 75 villages in Karnataka, India. In 2011, the authors conducted a survey in the 75 villages. The villages span 5 districts and range 2-3 hour drive outside of Bangalore. They are far enough apart to be treated as independent systems (the median distance is 46km and a district has between 1000-3000 villages). The survey included a village questionnaire, a census of all households, demographic covariates (including caste and occupation), as well as data on a number of amenities (e.g., roofing, latrine or electricity access quality). A detailed individual level survey was administered to most adults in every village. The survey included a networks module with twelve dimensions of relationships including financial relationships, social relationships, and advice relationships.\textsuperscript{30}

Our analysis focuses on two types of networks: the financial graph, $L_f$, and the social graph, $L_s$. The financial graphs represent risk-sharing connections and the social graph represents friendships/links used to socialize. We build “AND” networks, which say a link exists if it exists on every dimension being considered (various types of financial connections on the financial network, and various types of social connections on the social network).\textsuperscript{31}

The advantage of doing this is that it generates a network structure that is more robust to independent measurement error.\textsuperscript{32} In some of our empirical analysis, we will explicitly consider how our predictions differentially play out in $L_f$ relative to $L_s$, as our theory speaks to the former.

As our theory pertains to networks formed from multiple groups, here we make use of caste. Motivated by the social structure of our communities, and following Munshi (2006) and Banerjee et al. (2013), we partition our individuals into two caste groups: scheduled caste/scheduled tribes (SC/ST) and general merit/otherwise backward castes (GM/OBC). These are governmental designations used to condition the allocation of, for instance, school seating by caste and reflect a core fissure in the social fabric.

Table 1 provides some summary statistics for our data. The average number of households per village is 209 with a standard deviation of 80. The average degree of the financial graph is 3.1 and the social graph is 1.7. It is interesting to note that the financial graph is more dense. Further, we see that the clustering in the financial graph is 0.18 (0.06) whereas for the social graph it is 0.05 (0.03). Though the density is only 1.8 times higher for the financial graph than the social graph, the fact that the clustering is over 3 times higher for the financial graph is of note. This is consistent with the results of, for instance, Jackson et al. (2012) that financial links may need to be supported/embedded in cliques to sustain cooperation. Both the financial and social networks exhibit relatively few cross-caste links. This is seen looking at the ratio of the probability of having a cross-caste link relative to the probability of having a within-caste link. A within-caste link is five times more likely in the financial graph and three times more likely in the social graph. Finally, 67% of households are high caste (GM or OBC). We describe our main outcome variables later.

\textsuperscript{30}See Banerjee et al. (2014) for more details. In total we have network data from 89.14% of the 16,476 households based on interviews with 65% of all adult individuals aged 18-55.

\textsuperscript{31}We say $l_{ij} \in L_f$ if $i$ goes to $j$ to borrow money in times of need, $j$ goes to $i$ to borrow money in times of need, $i$ goes to $j$ to borrow material goods such as kerosene, rice or oil in times of need, and $j$ goes to $i$ to borrow material goods in times of need.

\textsuperscript{32}We say $l_{ij} \in L_s$ if $i$ goes to $j$’s house to socialize, $j$ goes to $i$’s house to socialize, $i$ goes to $j$ for advice, and $j$ goes to $i$ for advice.
Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Households per village</td>
<td>209.27</td>
<td>80.03</td>
</tr>
<tr>
<td>Average degree (financial network)</td>
<td>3.12</td>
<td>1.10</td>
</tr>
<tr>
<td>Average clustering (financial network)</td>
<td>0.18</td>
<td>0.07</td>
</tr>
<tr>
<td>Probability of cross-caste link/ Probability of within-caste link (financial network)</td>
<td>0.21</td>
<td>0.23</td>
</tr>
<tr>
<td>Average degree (social network)</td>
<td>1.75</td>
<td>0.77</td>
</tr>
<tr>
<td>Average clustering (social network)</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Probability of cross-caste link/ Probability of within-caste link (social network)</td>
<td>0.36</td>
<td>0.31</td>
</tr>
<tr>
<td>Fraction high caste (GM/OBC)</td>
<td>0.67</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Notes: For each variable we present the mean and the cross-village standard deviation. For the Myerson distances we use the approximation algorithm developed in the paper.

6.2 Empirical Framework

6.2.1 Predictions

The broad predictions of our model are that (1) there is endogenous centrality; (2) there is no underinvestment within groups; (3) agents cannot be too far away from each other (in terms of the Myerson distance); (4) nodes that have across caste links should be more Myerson central. However, these predictions are too general, in the cases of (2) and (3) they depend on an unobservable linking cost parameter, and there are many alternative stories, including ones not directly connected to risk-sharing, that generate similar predictions. For example, if individuals have heterogenous time budgets and made random links within and across groups, predictions (1) and (4) are mechanically generated.

For the above reasons, we need to turn to more subtle predictions that rely on the comparative statics under our model as we change parameters of the economic environment. We study four more demanding predictions from the theory with richer empirical content.

The first two predictions look at how network structure, described by Myerson distance, depends on income variability and correlation. The intuition is that in villages where the gains from risk sharing are higher (income is more variable/less highly correlated), the Myerson distance cannot be too high in equilibrium. Otherwise a pair of villagers would be incentivized to form an additional link. Doing so would enable them to appropriate a sufficiently large additional share of the gains from risk sharing that the social investment would be worthwhile. When the gains from risk sharing are lower, the Myerson distance can be larger.

The latter two predictions describe the composition of across-caste bridging links. Individuals with higher Myerson centrality have better incentives to form bridging links. However, when income variability or the within- versus across-caste income correlation is high, many members of either caste group can find it worthwhile to form a cross-caste bridging link. Thus, we expect the average centrality within their own group of forming the cross-caste link is lower when the incentives to form such links are higher. We now formalize our predictions.

In the case of one group, Proposition 7 provides the key characterization of the set of pairwise stable networks. This characterization yields an exact expression for increased payoffs two agents would receive were they to form an additional link. For a risk-sharing network to be stable, these benefits should be less than the cost of forming the link: for every $i$ and
that do not have a link,

\[ md(i, j, L) \leq \frac{2\kappa_w}{(1 - \rho_w)\lambda \sigma^2}. \]  

(5)

In our empirical setting, we are interested in Indian village networks where there are multiple groups, given by caste. Nevertheless, the inequality in (5) provides an appropriate benchmark for within-group links. Recall that the left side of the inequality captures the probability that the link is not essential for a random permutation in the arrival order and the right side gives the value of the variance reduction obtained. The key complication is that when a within-group link is essential for a subgraph, but it connects two otherwise separate components that contain people from multiple groups, then the value of the variance reduction will depend on the composition of people within the two components.

Consider the variance reduction obtained by combining any two components, with agents \( S_0, \ldots, S_k \) and \( \tilde{S}_0, \ldots, \tilde{S}_k \) respectively. From equation 2, this certainty equivalent value of this variance reduction is:

\[
[(1 - \rho_w) + f(s_0, \ldots, s_k, \hat{s}_0, \ldots, \hat{s}_k)(\rho_w - \rho_a)] \frac{\lambda \sigma^2}{2}.
\]

This term is hard to compute in general. Here we approximate it by assuming that the distribution of people from each group is the same within each component, so \( s_i = \alpha \) and \( \hat{s}_i = \beta \) for \( i = 0, \ldots, k \).

**Proposition 18** The certainty equivalent value of variance reduction obtained by linking a component with \( \alpha \) people from each of groups 0, \ldots, \( k \) to a component with \( \beta \) people from each of groups 0, \ldots, \( k \) is:

\[
\frac{(1 - \rho_w) \lambda \sigma^2}{2}.
\]

**Proof.** See the Appendix A. ■

Proposition 18 shows that the variance reduction obtained by permitting two components containing agents from multiple groups to risk share is the same as when all agents are from the same group, as long the proportion of people from each group is the same in each component.

The above considerations lead to the following predictions:

P1. In villages with higher \( \sigma^2 \), the average Myerson distance should be smaller.

P2. In villages with lower \( \rho_w \), the average Myerson distance should be smaller.

While our first set of predictions looks at the relationship between the Myerson distance and the environmental parameters, our second set of predictions looks at the composition of the links. Our interest is in which agents provide the across group links. Proposition 16 shows more central agents have better incentives to provide an across-group link. The importance of centrality will depend on the overall strength of the incentives to form across-caste links. When income variance is high, or within-caste income correlation is high relative to across-caste income correlation, the incentives to form an across-caste link will also be high and so network position will be less important; villagers in more varied locations will have sufficient incentives to form across-caste links. More formally, from the variance reduction given in equation 2 it is straightforward to show that for an across-caste bridging link \( l_{ij} \):
\[
\frac{\partial \Delta \text{Var}(L, L \cup \{l_{ij}\})}{\partial \sigma^2} > 0, \quad \frac{\partial \Delta \text{Var}(L, L \cup \{l_{ij}\})}{\partial \rho_w} > 0, \quad \frac{\partial \Delta \text{Var}(L, L \cup \{l_{ij}\})}{\partial \rho_a} < 0.
\]

This means that the incentives to form an across caste link are increasing in \(\sigma^2\) and \(\rho_w - \rho_a\), leading to the following predictions:

P3. In villages with higher \(\sigma^2\), the association between within caste centrality and providing across caste links is lower.

P4. In villages with higher \(\rho_w - \rho_a\), the association between within caste centrality and providing across caste links is lower.

6.2.2 Empirical Strategy

Our predictions are about how network structure varies with \(\sigma^2\), \(\rho_w\) and \(\rho_a\). The analysis is observational (not causal) and simply looks at the cross-sectional variation of network structure with these parameters through OLS.

We take two approaches to support our empirical claims. First, by focusing on different aspects of network structure (Myerson distance or the composition of cross-caste links), we provide evidence in support of our predictions from very different moments of the data.

Second, we exploit multigraph data. Our theory is built for risk-sharing networks and not for social connections. Exploiting this feature allows us to take a differences-in-difference approach and to study whether the correlations we document come from \(L^F\), the financial graph, and as opposed to \(L^S\), the social graph. Since the theory is differentially more informative about the structure of financial links as opposed to social links, the difference in patterns across link-types is informative.\(^{33}\) Furthermore, as unobserved endogeneity or homophily is likely to affect several dimensions of the multigraph at once, looking at the difference in risk-sharing link patterns versus social link patterns within village allows us to address unobserved village-level endogeneity that enters additively through a fixed effect. To take a simple example for a confounder, consider P1. In villages where the weather is more variable, less days are suitable for working so people may spend more time socializing, thereby spuriously generating shorter Myerson distances on average. Our difference-in-difference eliminates this sort of confound.

At the same time, we note that our empirical approach is more conservative than similar studies in the literature (e.g., Karlan et al. (2007), Ambrus et al. (2014), Kinnan and Townsend (2014)) in terms of statistical inference. First, these studies typically have very few networks (two, one, and 16, respectively), and therefore consider node or link-level regressions with standard errors generated at that level. This effectively treats nodes or dyads as independent or loosely correlated, making inference preclude, essentially, village-level shocks. Correlation in any factors outside the model (e.g., incentives to form links for other reasons) as well as equilibrium selection at the village level are all precluded from econometric analyses that don’t study the theory where the entire graph is the unit of observation. We make no such assumptions on the independence of nodes or dyads for valid statistical inference and instead exploit the fact that we have 75 independent villages (recall that the median pairwise distance is over 46km). By focusing on village-level variation, we are allowing for

\(^{33}\)A simple extension of the model delivers this prediction.
arbitrary correlation within graphs. In fact, in our most conservative specifications, we allow correlation at the subdistrict level.

Second, analyses in the previous literature usually do not have access to different types of edges – the multigraph – and therefore cannot employ our difference-in-differences approach. What we do, relative to this, is extremely conservative. By differencing across network-type, we are asking whether the patterns in the graph which match our theory are differentially at play for the financial network relative to the social networks.

Third, another reason this is conservative is because typical models of multigraph link formation have a fixed-cost component, so incentives driving risk-sharing links are likely to influence the structure of social or information links through this channel. The fixed cost component of the variation that is consistent with our theory is not being used in our analysis.

6.3 Variable Construction

6.3.1 Approximating the Myerson Distance and Centrality

Our next task is to compute the Myerson distance of every pair in every village and the Myerson centrality for all nodes. As it turns out, this is computationally infeasible for the sample sizes of our data (see Algaba, 2007). Thus, we develop an approximation, described below.

Let $md(L)$ be the matrix of Myerson distances. As before, $q(L) = 1/2 - md(L)$. So $q(L)$ is a matrix with the $ij$th entry capturing the probability that, upon being added to the network, agent $i$ will not be connected to agent $j$. It is difficult to directly characterize $md(L)$ (or equivalently, $q(L)$). Given that each village typically consists of around 230 households, there is an exponential (in the size of the network) number of candidate paths between each $i$ and $j$. Correctly accounting for paths that share nodes is computationally very intensive (see Proposition 7), and it has to be done for all unconnected pairs of agents. Instead, we develop a computationally feasible approximation of $md(L)$, which is exact for trees.

To approximate $q$, we use the following idea. The inclusion-exclusion principle combines paths in a way that avoids double counting. This involves two things. First, only minimal paths are included (paths that are a strict superset of other paths are excluded). Second, it provides the right way to add and subtract minimal paths that share nodes. This is complicated and we abstract from it. However, we will be able to include only minimal paths and weight these paths appropriately. We denote our approximation of $q$ by $\hat{q}$.

More precisely, we start with a node, move to its neighbors, then move to its neighbors’ neighbors, and so on, all the while never returning to a previously used node along a given path. This allows us to avoid counting paths that are not minimal. All the while, we keep track of how many ways we have moved from the original node to any given node. This gives us a count of the minimal paths. With these minimal paths in hand, we can add them up weighting a path of length $l$ by $1/l$. By the inclusion-exclusion principle, this is correct weighting for independent paths.

Figure 3 presents two examples: a tree and a circle. The tree has a single minimal path between nodes 1 and 8, whereas the circle has two between nodes 1 and 4. Figure 4 shows how links are removed for the case of a tree. After a node is used, links from that node to its neighbors are deleted, though links to that node are still allowed.

Further, due to presumed measurement error (see Banerjee et al. (2013)), there are likely to be missing paths. In fact, the data have occasional disconnected components and thus measures that are precisely based on exact paths or even maximal path lengths are likely to be problematic (Chandrasekhar and Lewis (2014)).
Figure 3: The $i,j$ – when we are computing $md(i, j, L)$ – are given in purple. The tree contains a single minimal path (orange) whereas the circle contains two (orange and blue).

Figure 4: As the algorithm progresses directed links to nodes that have already been reached are deleted. This makes sure only minimal paths are included. In this case, as in all tree networks, there is a unique minimal path from A to B.

To build intuition and show our algorithm’s limitations, consider a circle with 7 nodes. We want to compute $md(1, 4)$ as in Figure 3. We need to compute the quantity for each minimal path, but by the inclusion-exclusion principle, need to subtract the event that both are present. The long path contributes $1/5$ (1 added last in $\{1,7,6,5,4\}$), the short path $1/4$, and by the inclusion-exclusion principle we subtract $1/7$ (1 added last in $\{1,2,...,6,7\}$). Thus we see

$$q(1, 4) = \frac{1}{5} + \frac{1}{4} - \frac{1}{7}$$

Algorithm 19 (Approximation of $q$) Let $e^i$ be the $i$th basis vector. Initialize $\hat{q} = \text{zeros}(n, n)$, a matrix of zeros. Initialize $z^{t,i} = \text{zeros}(n, 1)$ and $x^{t,i} = \text{zeros}(n, 1)$ to be $n$-vectors of zeros, indexed by $i = 1,...,n$ and $t = 1,...,T$. Repeat steps (1-4) for all $(e^1,...,e^n)$ to compute $\hat{q}$.

1. Period 1: There is no identification nor updating steps.

   (a) Percolation: $x^{1,i} = Ae^i$.

2. Period 2, given $(x^{1,i}, A)$:
(a) Identification: $z^{2,i} = e^i$.

(b) Update graph: $A_2 = \text{zeros}(n,n)$, $A_2(:, -z^{2,i}) = A(:, -z^{2,i})$.

(c) Percolation: $x^{2,i} = A_2 x^{1,i}$.

3. Period $t$, given $(x^{t-1,i}, A_{t-1})$:

(a) Identification: $z^{t,i} = 1\left\{\sum_{s=2}^{t} x^{s-2,i} > 0\right\}$.

(b) Update graph: $A_t = \text{zeros}(n,n)$, $A_t(:, -z^{t,i}) = A_{t-1}(:, -z^{t,i})$.

(c) Percolation: $x^{t,i} = A_t x^{t-1,i}$.

4. Set $\hat{q}_{ij} = \sum_{t=1}^{T} x_{j}^{t,i} \left(\frac{1}{t+1}\right)$ where $x_{j}^{t,i}$ is the $j$th entry of $x^{t,i}$, but $\hat{q}_{ij} = 0$ if $l_{ij} \in L$.

**Proposition 20** Let $L$ be a connected tree. Then $\hat{q}(L) = q(L)$.

**Proof.** See the Appendix A. ■

Note that $\hat{q}(L)$ works by treating all minimal paths as independent while some rely on the same nodes. In general each $\hat{q}_{ij}$ weakly overestimates $q_{ij}$. To operationalize this in our regression analysis, we need a village level measure of Myerson distances. We use $\tilde{q}(L) := \sum_{i<j} \hat{q}_{ij} / \binom{n}{2}$ which measures an appropriately weighted density of the network. Finally, to approximate Myerson centrality we use $\sum_j \hat{q}_{ij}$ as people are central when they are likely to be connected to others. Thus, their $q_i$ terms are high, or alternatively, they have low Myerson distance to others.

### 6.3.2 Construction of Income Variability and Caste-Income Correlation

Our analysis requires estimates of $\sigma^2$ and $(\rho_w - \rho_a)$. However, due to data limitations – despite extremely detailed multigraph network data across many villages, we lack financial records – we need to construct measures of these quantities.

For income variability, we merge National Oceanic and Atmospheric Administration (NOAA) data with our network data. Matsuura and Willmott (2012) construct a gridded monthly time series of terrestrial precipitation from 1900-2010. We match this to our villages using our GPS data and the crucial variable is the standard deviation of rainfall by village once removing month fixed-effects.

Measuring income correlation is difficult. Ideally, we would have time series data on incomes of all households, as well as plausible instruments, allowing us to calculate the exogenous variation in income correlation both within and across caste for each village. While we do not have access to income data, we do have detailed data on occupation.

Thus, we make use of the relative within-caste to across-caste occupation correlation. The main idea is that shocks to individuals will be more correlated when they have the same occupation. We take two approaches to computing caste-income correlation.

---

35We recognize that occupations certainly have a choice-component. Nevertheless we proceed with these measures for three reasons. First, in rural villages the primary household occupation (agriculture or sericulture) is often passed on through generations. Second, it is possible to show that under a natural model with endogenous selection into occupation, the within versus across group income correlations are captured by our occupational choice measures, whereas the choice of occupation do not generate spurious correlations with the network in a manner consistent with P1-P4. Finally, this is the best approximation given the severe income data limitations, as a necessary component for our analysis is the network data.
One approach is to look at the correlation of being in the high-caste group with holding a given occupation, for all occupations in our survey. We then take the weighted-average of these correlations, where the weight is by the share of agents in the occupation. Thus, for a given village we consider:

\[
\hat{\rho}_w - \hat{\rho}_a := \sum_{k=1}^{K} \text{corr}(\text{Caste}, \text{Occupation}_k) \cdot P(\text{Occupation}_k),
\]

where \( \text{Caste} \) is an \( n \times 1 \) vector of GM/OBC dummies and \( \text{Occupation}_k \) is an \( n \times 1 \) vector of dummies for a household having a member in occupation \( k \). This constructs a score which is 0 if there is no correlation between caste group and occupation and 1 if caste group perfectly predicts occupational choices.

Another approach involves making a few simplifying assumptions about the structure of the income process. If the income of an individual in a given caste and occupation can be thought of as depending on a caste-occupation specific mean, occupation-specific idiosyncratic iid shocks, as well as individual-specific idiosyncratic iid shocks, where the occupation shocks all have the same variance across occupations, then as shown in Appendix B, we can write

\[
\hat{\rho}_w - \hat{\rho}_a = \sum_{g} \phi_g p(o_{i,g} = o_{j,g}) - P(o_{i,A} = o_{j,B}),
\]

where \( o_{i,g} \) denotes the occupation of \( i \) in caste \( g \) and \( \phi_g \) is the population share of caste \( g \). In sum, both measures are intuitive, but imperfect, proxies for the regressors we need in our analysis. Therefore we take a rough-and-ready approach, utilizing both in our analysis.

1. \( \hat{\rho}_w - \hat{\rho}_a I = \sum_{g} \phi_g p(o_{i,g} = o_{j,g}) - P(o_{i,A} = o_{j,B}). \)
2. \( \hat{\rho}_w - \hat{\rho}_a II = \sum_{k=1}^{K} \text{corr}(\text{Caste}, \text{Occupation}_k) \cdot P(\text{Occupation}_k). \)

### 6.4 Results

#### 6.4.1 Myerson Distance as a Function of Income Variability and Correlation

We begin with P1 and P2. Figure 5 presents results in the raw data. As predicted, villages in areas corresponding to more variable income processes are associated with lower \( \sim \text{md}(L^F) \). Villages with more within-caste income correlation are associated with higher \( \sim \text{md}(L^F) \).

Table 2 demonstrates the robustness of this graphical evidence in regression of \( \sim \text{md}(L^F) \) on \( \sigma, \rho_w, \) and with district or subdistrict fixed effects as well as controls for caste composition.\(^{36}\) In all specifications here and throughout, unless otherwise noted we use a Wild clustered bootstrap to account for subdistrict level clustering in our inference (Cameron et al., 2008).\(^{37}\) Columns 1-3 present regressions using \( \hat{\rho}_w I \) whereas columns 4-6 present regressions using \( \hat{\rho}_w - \hat{\rho}_a II \). A one standard deviation increase in Measure I is associated with a 0.296 standard deviation increase in \( \sim \text{md}(L^F) \) (column 1). However, there is not a significant relationship

---

\(^{36}\)Similar results using \( 1/D(L^F) \), the inverse of average degree, are in Appendix C.

\(^{37}\)While these villages are essentially independent units, the median distance being 46km apart, given that geography is a determinant of rainfall variability and occupation, we take a conservative approach. Our villagers are members of 12 subdistricts and we therefore cluster our standard errors at that level. To deal with the finite sample bias we use a Wild cluster bootstrap procedure for the \( t \)-test statistic, using Rademacher weights, to generate \( p \) values for hypothesis testing as well as standard errors.
between income variability and $\tilde{md}(L^F)$ in this specification. Column 4 presents the analogous results using $\hat{\rho}_w - \rho_{a}^{II}$. Again a one standard deviation in Measure II is associated with a 0.205 standard deviation increase in $\tilde{md}(L^F)$, but income variability has no detectable correlation with the outcome variable (column 4). In columns 2 and 5 we add district fixed effects and the results are mostly reflective of those in columns 1 and 3. A one standard deviation increase in income correlation corresponds to a 0.11 or 0.18 standard deviation increase in Myerson distance, using Measure I ($p$-value 0.12) and Measure II ($p$-value 0.012) respectively. Finally, when we include subdistrict fixed effects, our estimate under Measure I is too noisy to distinguish from zero, though we find similar evidence under Measure II (a 0.1 standard deviation effect, $p$-value 0.13). Further, when comparing villages within subdistrict we are able to identify an income variability effect. A one standard deviation increase in income variability is associated with a 0.28 or 0.3 standard deviation decrease in the average Myerson distance (columns 3 and 6, respectively).

Taken together, we show increases in within-caste income correlation is associated with
Myerson distance in a manner consistent with our theory though the evidence for income variability is considerably weaker.

6.4.2 Differences by network type

Next we use a difference-in-differences approach and see whether the effects we are interested are coming differentially from the financial graph as opposed to the social graph. Figure 6 presents the raw data, differenced, graphically. Villages in areas corresponding to more correlated within-caste income processes are associated with greater \( \tilde{md}(L^F) - \tilde{md}(L^S) \). Similarly, an increase in the income variability is associated with a differential decrease in the measure of network density in the financial graph as compared to the social graph.

![Figure 6: Myerson distance, differenced across network type, versus income variability or within-caste income correlation.](image)

Figure 6: Myerson distance, differenced across network type, versus income variability or within-caste income correlation.

Let \( v \) index village and \( t \in \{F,S\} \) index network type. We use the following regressions:

\[
Y_{v,t} = \alpha + \beta \cdot \sigma_v \cdot 1_{\{t=F\}} + \gamma \cdot \rho_{w,v} \cdot 1_{\{t=F\}} + \mu_v + \delta \cdot X_v \cdot 1_{\{t=F\}} + \epsilon_{v,t},
\]

where \( Y_{v,t} = \tilde{md}(L^t_v) \), \( \mu_v \) is a possibly endogenous village fixed-effect, and \( X_v \) are village-level controls. P1 and P2 correspond to \( \beta < 0 \) and \( \gamma > 0 \), since our theory pertains only to risk-sharing networks. Thus, our effects should be differentially more predictive for the financial network and not the social network and should remain true when looking at relative effects.

Table 3 presents the results. We find that a one standard deviation increase in income variability differentially decreases \( \tilde{md}(L) \) by about 0.24 standard deviations more in financial networks than social networks (columns 1, 4), irrespective of the measure of income correlation used. Similarly, a one standard deviation increase in Measure II differentially increases \( \tilde{md}(L) \) by 0.16 standard deviations more in financial networks than social networks (column 2). However, we find no association using Measure I for the income correlation.

Thus, even when we look within-village, by allowing for village fixed effects, and yet allowing for subdistrict level correlation in our error terms, we see that the financial graph behave in a manner consistent with the theory relative to the social graph.

6.4.3 Association between within-caste centrality and cross-caste links

We now look at how the composition of cross caste links vary with these parameters. P3 shows that in villages with higher variability we should see a greater association between
Table 3: Myerson distance vs. income variability and correlation, with village FE

<table>
<thead>
<tr>
<th></th>
<th>Measure I</th>
<th>Measure II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Income Correlation x 1{Financial network}</td>
<td>0.013</td>
<td>0.162</td>
</tr>
<tr>
<td></td>
<td>(0.088)</td>
<td>(0.084)</td>
</tr>
<tr>
<td></td>
<td>[0.918]</td>
<td>[0.058]</td>
</tr>
<tr>
<td>Income Variability x 1{Financial network}</td>
<td>-0.245</td>
<td>-0.234</td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.131)</td>
</tr>
<tr>
<td></td>
<td>[0.098]</td>
<td>[0.096]</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.8100</td>
<td>0.8170</td>
</tr>
</tbody>
</table>

Note: Outcome variable is the average Myerson centrality of the graph of a given type (financial or social). Outcome variable, income correlation, and income variability all scaled by their standard deviations. All regressions include village fixed effects. Specifications include control for caste composition: p(1-p), where p is share in high caste, interacted with network type. Wild clustered bootstrap standard errors are presented in (), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild clustered bootstrap t are presented in [.].
Table 4: Association between average within-caste centrality of nodes with cross-caste links and measures of rainfall variability and within-caste income correlation

<table>
<thead>
<tr>
<th></th>
<th>Myerson Centrality</th>
<th>Eigenvector Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Income Correlation x 1{Financial network}</td>
<td>-0.222</td>
<td>-0.252</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.036)</td>
</tr>
<tr>
<td></td>
<td>-0.112</td>
<td>-0.148</td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.060)</td>
</tr>
<tr>
<td></td>
<td>-0.191</td>
<td>-0.064</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Income Variability x 1{Financial network}</td>
<td>0.261</td>
<td>-0.313</td>
</tr>
<tr>
<td></td>
<td>(0.201)</td>
<td>(0.295)</td>
</tr>
<tr>
<td></td>
<td>-0.130</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td>(0.129)</td>
</tr>
<tr>
<td></td>
<td>-0.261</td>
<td>-0.291</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td>(0.170)</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>District x 1{Financial network} FE</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Subdistrict x 1{Financial network} FE</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.196</td>
<td>0.529</td>
</tr>
<tr>
<td></td>
<td>0.675</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>0.137</td>
<td>0.346</td>
</tr>
</tbody>
</table>

Notes: Outcome variable is the Myerson centrality, computed using our approximation algorithm, or eigenvector centrality. Income correlation given by Measure I. Outcome variables and regressors are scaled by their standard deviations. All specifications use village fixed effects and caste composition controls. Columns 2 and 4 include district-by-network type fixed effects and columns 3 and 6 include subdistrict-by-network type fixed effects. Wild clustered bootstrap standard errors are presented in ( ), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild clustered bootstrap t are presented in [ ].

6.5 Discussion

This section provided suggestive evidence, consistent with our theory, demonstrating:

1. Networks have higher Myerson distance with more within group income correlation.
2. Networks exhibit lower Myerson distance when there is more variable rainfall.
3. Networks exhibit stronger associations between within-group centrality and across-group linking when within-group income correlation is higher, though we detect no relationship with variability of rainfall.
4. Results (1)-(3) are robust to a difference-in-differences approach. By differencing out the social graph, we can remove village-fixed endogenous factors and more convincingly argue that the underlying force concerns risk sharing.

Having several distinct prediction is useful as we are conducting an observational analysis. That all significant correlations are consistent with our theory, and we find no falsification, is encouraging.

Of course, there are several other stories consistent with the income variability prediction. For instance, one may think that any theory where the returns to investing in risk-sharing relationships go up should predict more dense networks. We note, however, that this prediction does not hold in Bramoulle and Kranton (2007a), which is perhaps the closest paper to ours. In their model, increased variability has no effect on the equilibrium network structure.

Furthermore, our compositional story is new. More generally, P3-P4 are either specific to the risk sharing relationships between castes or differences in income correlations between castes, and so within the literature are unique to our model. These predictions are also more subtle. For instance, higher within group correlation leading to more sparse risk sharing networks is not a conclusion of other risk-sharing network formation stories.
We again emphasize that our findings are clearly observational (non-causal) and subject to measurement error. However, given the unique data – 75 independent village network data and multigraph data allowing us to difference out by link-type to see if effects are driven by those consistent with our theory – this represents a first opportunity to tackle the type of questions we address. The data requirements are immense, and we are able to take a serious pass at looking at different types of cuts of our data – several network measures, several parameters of the economic environment, and several different results from our theory – to see if the data is consistent with our story. Additionally, we are extremely conservative in conducting statistical inference, by only relying on the independence across-villages (in fact we are more conservative, employing a Wild cluster bootstrap at the subdistrict level) and not exploiting any of the within-village observations (which are likely to be correlated).

7 Conclusion

In this paper we develop a relatively tractable model of network formation and surplus division in a context of risk-sharing, that allows for heterogeneity in correlations between the incomes of pairs of agents. Such correlations have a big impact on the potential of informal risk sharing to smooth incomes. We investigate the incentives for relationships that enable risk sharing to be formed both within a group (caste or village) and across groups giving access to less correlated income streams.

We find that overinvestment into social relations is likely within a group, but there is potential underinvestment into more costly social connections that bridge different groups. We also find a novel trade-off between equality and efficiency. Within groups the social structure that has the highest level of inequality is the star. The efficient network that minimizes the incentives to overinvest within group and maximizes the incentives to establish across group links is also the star. More generally, having agents located centrally within their group improves the incentives for across group links to be established and reduces the incentives of others to form superfluous within group links that redistribute but do not create surplus.

Using a unique dataset of 75 Indian villages, we find empirical support for our model. This is particularly important because our model abstracts from the more widely studied enforcement role it might play. Although we certainly do not reject the network also playing such a role, we see evidence in our data that the network structure affects how the gains from risk sharing are distributed, suggesting that this is an important consideration when studying the network formation problem. In particular, we find that higher income variability is associated with denser social network (in a sense formally defined by the model), and that more centrally connected individuals are more likely to establish across group links, and this association becomes stronger when within group income correlation increases relative to across group income correlation. Moreover, we find evidence that these relationships are differentially stronger for the network of financial relationships than for the network of other types of social relationships.

Although we focus our analysis on risk-sharing, our conclusions regarding network formation can apply in other social contexts too, as long as the economic benefits created by the social network are distributed in a similar way to our model - a question that requires further empirical investigation. There are many directions for future research in the theoretical analysis as well, even within the context of risk-sharing, a natural next step would be to provide a dynamic extension of the analysis that allows for autocorrelation between income
realizations as well.

References


A Supplementary Proofs

Proof of Proposition 1. To prove the first statement, consider villagers’ certainty equivalent consumption. Let $\hat{K}$ be some constant and consider the certain transfer $K'$ (made in all states of the world) that $i$ requires to compensate him for keeping a stochastic consumption plan $c_i$ instead of another stochastic consumption plan $c'_i$:

\begin{align*}
E[u(c_i + \hat{K})] &= E[u(c'_i + \hat{K} - K')] \\
-\frac{1}{\lambda} e^{-\lambda \hat{K}} E[e^{-\lambda c_i}] &= \frac{1}{\lambda} e^{-\lambda \hat{K}} e^{\lambda K'} E[e^{-\lambda c'_i}] \\
e^{\lambda K'} &= \frac{E[e^{-\lambda c_i}]}{E[e^{-\lambda c'_i}]} \\
K' &= \frac{1}{\lambda} \left( \ln \left( \frac{E[e^{-\lambda c_i}]}{E[e^{-\lambda c'_i}]} \right) - \ln \left( E[e^{-\lambda c'_i}] \right) \right)
\end{align*}

This shows that the amount $K'$ needed to compensate $i$ from keeping the more stochastic consumption stream $c_i$ instead of consumption stream $c'_i$ is independent of $\hat{K}$. As villagers’ certainty equivalent consumption for a lottery is independent of his consumption level, certainty equivalent units can be transferred among the villagers without affecting their risk preferences, and expected utility is transferable.

Next we exactly characterize the set of Pareto efficient risk sharing agreements. Borch (62) and Wilson (68) showed that a necessary and sufficient condition for a risk sharing arrangement between $i$ and $j$ to be Pareto efficient is that in all states of the world $\omega \in \Omega$:

\begin{align*}
\frac{\partial u_i(c_i(\omega))}{\partial c_i(\omega)} &= \frac{\partial u_j(c_j(\omega))}{\partial c_j(\omega)} = \alpha_{ij}
\end{align*}

where $\alpha_{ij}$ is a constant. Substituting in the CARA utility functions, this implies that:

\begin{align*}
e^{-\lambda c_i(\omega)} &= \alpha_{ij} \\
c_i(\omega) - c_j(\omega) &= -\frac{\ln(\alpha_{ij})}{\lambda} \\
E[c_i(\omega)] - E[c_j(\omega)] &= -\frac{\ln(\alpha_{ij})}{\lambda} \\
c_i(\omega) - c_j(\omega) &= E[c_i(\omega)] - E[c_j(\omega)] \quad (6)
\end{align*}

Letting $i$ and $j$ be neighbors, such that $j \in N(i)$, equation 6 means that when $i$ and $j$ reach any Pareto efficient risk sharing arrangement their consumptions will differ by the same constant in all states of the world. Moreover, by induction that same must be true for all path connected villagers.

Consider now the problem of splitting the incomes of a set of villagers $S$ in each state of the world to minimize the sum of their consumption variances:

\[
\min_c \sum_{i \in S} Var(c_i)
\]

subject to $\sum_{i \in S} y_i(\omega) = \sum_{i \in S} c_i(\omega)$ in all possible states of the world $\omega$. Note,
\[ \sum_{i \in S} \text{Var}(c_i) = \sum_{i \in S} \sum_{\omega \in \Omega} p(\omega)(c_i(\omega) - E[c_i])^2 \]

where \( p(\omega) \) is the probability of state \( \omega \). As the sum of variances is convex in consumptions and the constraint set is linear the maximization is a convex program. The first order conditions of the Lagrangian are that for each \( i \in S \) and each \( \omega \in \Omega \):

\[ 2(c_i(\omega) - E[c_i]) = \gamma(\omega), \]

where \( \gamma(\omega) \) is the Lagrange multiplier for the state \( \omega \). Thus:

\[ c_i(\omega) - c_j(\omega) = E[c_i(\omega)] - E[c_j(\omega)] \]

for all \( i, j \in S \). This is exactly the same condition as the necessary and sufficient condition for an ex ante Pareto efficiency. Hence, a risk sharing agreements is Pareto efficient if and only if the sum over all path connected villagers of consumption variances is minimized.

Using the necessary and sufficient condition for efficient risk-sharing, we obtain:

\[ \sum_{k \in S} y_k(\omega) = \sum_{k \in S} c_k(\omega) = |S||c_i(\omega) - \sum_{k \in S} (E[c_i(\omega)] - E[c_k(\omega)])| \]

\[ c_i(\omega) = \frac{1}{|S|} \sum_{k \in S} y_k(\omega) + \frac{1}{|S|} \sum_{k \in S} (E[c_i(\omega)] - E[c_j(\omega)]) = \frac{1}{|S|} \sum_{k \in S} y_k(\omega) + \tau_i, \]

where \( \tau_i = E[c_i(\omega)] - E[\sum_{k \in S} y_k(\omega)]. \)

**Proof of Proposition 7.** Agent \( i \) will have a net positive benefit from forming a link \( l_{ij} \) if and only if \( MV_i(L) - MV_i(L/l_{ij}) > \kappa_w \). We simply need to show that

\[ MV_i(L) - MV_i(L/l_{ij}) = MV_j(L) - MV_j(L/l_{ij}) = md(i, j, L)V. \]

Some additional notation will be helpful. Suppose agents are added to the network in a order determined by a permutation of the agents drawn uniformly at random. The random variable \( \hat{S}, \subseteq \mathbb{N} \) identifies the set of agents, including \( i \), who are drawn weakly before \( i \) in the permutation. We let \( L(\i, L) \) be the probability distribution over the networks \( L(\hat{S}_j) \), generated by a permutation being drawn uniformly at random. Finally, let \( q(i, j, L) \) be the probability that \( i \) and \( j \) are path connected on a network \( L(\hat{S}) \) drawn from \( L(\i, L) \).

The certainty equivalent value of the reduction in variance due to a link \( l_{ij} \) in a graph \( L' \), is \( V \) if the link is essential and 0 otherwise. The change in \( j \)'s Mysisen Value, \( MV_i(L) - MV_i(L/l_{ij}) \), is then just the product this probability and the value \( V \). This probability is the probability that \( j \) has already been added \((1/2)\) and there is no other path from \( i \) to \( j \). As \( i \) must have been added for there to be another path between \( i \) and \( j \), the probability that \( l_{ij} \) is essential is \( 1 - 1/2 - q(i, j, L) \). We complete the proof by establishing that

\[ q(i, j, L) = \sum_{k=1}^{P_{ij}(i, j, L)} (-1)^{k+1} \left( \sum_{1 \leq i_1 < \cdots < i_k \leq |P(i, j, L)|} \frac{1}{P_1 \cup \cdots \cup P_k} \right), \]

so \( 1 - 1/2 - q(i, j, L) = 1/2 - q(i, j, L) = md(i, j, L) \).

We can represent \( q(i, j, L) \) in terms of the paths in \( L \). First note that there is a path from \( i \) to \( j \) on \( L' \), if only if \( P(i, j, L') \neq \emptyset \); A non-minimal path exists only if a minimal
path exists. We therefore need to find the probability that there is at least one minimal path \( P_k(i, j, L) \in P(i, j, L) \) present when \( j \) is added. Let \( Pr(P_k(i, j, L)) \) be the probability of the event that agent \( j \) is the agent drawn last from the set \( P_k(i, j, L) \), in a random draw from \( L(j, L) \). This is a necessary and sufficient condition for the path \( P_k(i, j, L) \) to exist when \( j \) is drawn. The probability of this event is \( 1/|P_k(i, j, L)| \).

We need to find the probability that any path between \( i \) and \( j \) exists in a random draw from \( L(j, L) \), i.e., \( \cup_{P_k(i, j, L)\in P(i, j, L)} Pr(P_k(i, j, L)) \). As these paths are not disjoint, the inclusion-exclusion principle needs to be applied. Doing so results in the formula shown.\(^{39}\)

**Proof of Proposition 16.** Consider the set of arrival order permutations for all agents for the network \( L \cup l'_{ij} \). We will show that we can match each permutation in this set to a permutation for the arrival orders of the agents on the network \( L \cup l_{ij} \) such that: (i) \( i \) is added at the arrival time of \( i' \) in the original permutation; (ii) when added \( i \) connects to exactly the same set of agents \( N \setminus S_0 \) as \( i' \) connects to; (iii) \( i \) is connected to weakly more agents within \( S_0 \) when added than \( i' \). Equation 4 shows that the risk reduction, and hence the additional payoff to \( k \in S_0 \), from the across-group link \( l_{kj} \) is an increasing function of the component size of \( k \)'s groups. It then follows that:

\[
MV(i; L \cup l_{ij}) - MV(i; L) > MV(i'; L \cup l'_{ij}) - MV(i'; L).
\]

To match permutations we do the following. Take the set of original permutations for \( L \cup l'_{ij} \) and make the following adjustments. First, switch the arrival positions of \( i' \) and \( i \). This alone is enough to ensure that condition (i) and (ii) are satisfied. Now note that as \( i \) is more central than \( i' \) the CDF over agents \( i \) is connected to within \( S_0 \) upon being added first order stochastically dominates the set of agents that \( i' \) is connected to within \( S_0 \). Label the permutations for \( i'' \) of agents in \( S_0 \) in the following way. First consider all permutations in which \( i' \) is added first. Label these permutations \( P_{i'}(1) \ldots \). Now consider all permutations in which one other agent is added before \( i' \) and continue to label permutations by ordering these permutation in terms in ascending order of how many agents \( i' \) is connected to upon being added. Repeat until all permutations are ordered. Do the same for \( i \). Now make the following final adjustment to \( i' \)'s permutation over all agents. If the sub-permutation of agents within \( S_0 \) for \( i' \) is permutation \( P_{i'}(x) \), reorder the arrival of agents within \( S_0 \) to be \( P_{i}(x) \). This permutation for \( i \) now also satisfies condition (iii). By construction, \( i' \) will be connected to more agents within \( S_0 \) than \( i \) upon being added. The only remaining thing to check is that is that the matching of permutations is valid, and as each constructed permutation for \( i \) is unique the matching is valid.\(\)

**Proof of Proposition 18.** We simply substitute \( s_i = \alpha \) and \( \hat{s}_i = \beta \) for \( i = 0, \ldots, k \) into equation 2. This yields:

\[
\Delta \text{Var}(L \cup l_{ij}, L) = \left(1 - \rho_w\right) + \frac{2(k+1)^3\beta^2\alpha^2 - 2(k+1)^3\alpha^2\beta^2}{\left(\sum_{i=0}^{k} s_i\right)\left(\sum_{i=0}^{k} \hat{s}_i\right)\left(\sum_{i=0}^{k} s_i + \hat{s}_i\right)}(\rho_w - \rho_a) \right] \sigma^2
\]

\[
= (1 - \rho_w)\sigma^2.
\]

\(^{39}\)For example, if there are \( K \) nodes in \( P_k(i, j, L) \) and \( K' \) nodes in \( P_{k'}(i, j, L) \) then \( Pr(P_k(i, j, L)) = (1/K) \) and \( Pr(P_{k'}(i, j, L)) = (1/K') \). Similarly, if there are \( K'' \) distinct nodes in \( P_k(i, j, L) \cup P_{k'}(i, j, L) \) then the probability both paths exists is \( (1/K'') \). So, the probability that either path \( k \) or \( k' \) is present is the probability path \( k \) is present plus the probability path \( k' \) is present, less the probability that both are present: \( (1/K) + (1/K') - (1/K'') \).
Multiplying by $\lambda/2$ to get the certainty equivalent value of the variance reduction completes the proof. ■

**Proof of Proposition 20.** We will say that agent $k$ is a distance $t$ neighbor of $i$ if all minimal paths (i.e., paths that are not a superset of some other path) from $i$ to $k$ take exactly $t$ steps. As $L$ is a connected tree, there is a unique path from $i$ to $j$ and we can therefore assign each agent $k \neq i$ to be a distance $t$ neighbor, for some $t$.

Consider the implementation of the algorithm to find $\hat{q}_{ij}$. We begin by calculating $x_{1,i} = Ae_{i}$, where $e_{i}$ be the $i$th basis vector. This identifies all agents connected to $i$. We then delete the $i$th column from the adjacency matrix $A$ and set this new matrix equal to $A_2$. This deletes the outward links from $i$ in the network $L$. Starting from $i$'s neighbors we then find their neighbors on $A_2$. In other words we calculate $x_{2,i} = A_2 x_{1,i}$. This identifies the distance 2 neighbors of $i$. We then delete the columns of $A_2$ that are indexed by one of $i$'s neighbors and so on.

In the $t$th round the algorithm identifies the distance $t$ neighbors of $i$. Thus, for all $t < l$, $x_{j}^{l,t} = 0$, for $t = l$, $x_{j}^{l,1} = 1$ and for all $t > l$, $x_{j}^{l,i} = 0$. Thus, if the unique path from $i$ to $j$ is length $l$, $\hat{q}_{ij} = 1/(l + 1)$. From equation 2 it is also easily checked that $q_{ij} = 1/(l + 1)$. ■

**B Income Correlation from Occupation Correlations**

Here we present the rationale for our two caste income correlation measures. Because we lack income data, we must use occupation data in order to construct proxies. In our surveys we have occupation data for all surveyed individuals, coded as small business owner, land-owning farmer, farm laborer, dairy producer/cattle rearer, sericulture owner, sericulture laborer, government official, garment worker, industrial factory worker, industrialist, mason/construction worker, street vendor, artist (e.g., sculptor), and domestic help.

Let $y_{i,g}$ be the income and $o_{i,g} \in O$ the occupation of person $i$ in group $g \in \{A,B\}$. Also denote the probability that person $i$ is in occupation $o$ and $j$ is in occupation $o'$ by $p(o_{i,A} = o, o_{j,B} = o')$, where $g(i) = A$ and $g(j) = B$. Finally, it will be useful to denote by $\phi_g$ the proportion of individuals in group $g$.

In order to operationalize the quantity $\rho_w - \rho_a$ in our empirical exercises, we use

1. $\hat{\rho}_w - \rho_a = \sum_g \phi_g p(o_{i,g} = o_{j,g}) - p(o_{i,A} = o_{j,B})$ and
2. $\hat{\rho}_w - \rho_a = \sum_{k=1}^{K} \text{corr}(\text{Caste}, \text{Occupation}_k) \cdot P(\text{Occupation}_k)$.

The first measure looks at the difference in the (weighted) share of pairs in the same caste who hold the same occupation relative to the share of pairs in different castes who hold the same occupation. We show that this measure is exact when each individual draws an income independently which can have a mean that depends on her caste and occupation, there is an occupation level shock, but the occupation level shock has the same variance for every occupation.

The second measure presents a score that ranges from 0 to 1. If caste fully explains occupation (where there is therefore scope for maximal within caste income correlation), the score is 1. However, if caste does not explain occupation at all, the score is 0. This measure is computed as the average caste-occupation correlation, averaged over the occupations.

In what follows, assume the following:

$$y_{i,g} = \mu_{g,o} + \epsilon_o + u_i,$$
where $\epsilon_o$ is a mean zero, variance $\sigma_o^2$ iid shock that hits each occupation, $u_i$ is an iid shock with mean zero and variance $\sigma_u^2$ that hits each individual, and $\mu_{g,o}$ is a caste-occupation specific mean.

By the law of total covariance, we have:

$$
\text{cov}(y_{i,A}, y_{j,B}) = \sum_{o \in O} \sum_{o' \in O} \text{cov}(y_{i,A}, y_{j,B}|o_{i,A}, o_{j,B})p(o_{i,A} = o, o_{j,B} = o') + \text{cov}(E[y_{i,A}|o_{i,A}], E[y_{j,B}|o_{j,B}]),
$$

which we will use in our computations below.

### B.1 Identical means and variances

We begin with the simple case where all means and variances are identical. This justifies the use of $\rho_w = \rho_u^I$.

**Lemma B.1** Assume that $\mu_g = \mu, \forall g$ and $\sigma_o^2 = \sigma^2, \forall o \in O$. Then

1. $\rho_w \propto \sum_g \phi_g p(o_{i,g} = o_{j,g})$ and
2. $\rho_u \propto p(o_{i,g} = o_{j,g'})$,

both with the same constant of proportionality $\frac{\sigma^2}{\sigma^2 + \sigma_u^2}$.

**Proof.** It is immediately clear that

$$\text{cov}(y_{i,A}, y_{j,B}|o_{i,A} = o, o_{j,B} = o') = 0,$$

if $o \neq o'$, and if $o = o'$ then

$$\text{cov}(y_{i,A}, y_{j,B}|o_{i,A} = o, o_{j,B} = o) = E[(\epsilon_o)(\epsilon_o)] = \sigma^2.$$

Also, note that

$$\text{cov}(E[y_{i,A}|o_{i,A}], E[y_{j,B}|o_{j,B}]) = 0,$$

implying

$$\text{cov}(y_{i,A}, y_{j,B}) = p(o_{i,A} = o_{j,B})\sigma^2.$$

Thus the correlation between $y_{i,g}$ and $y_{j,g'}$ is

$$\text{corr}(y_{i,g}, y_{j,g'}) = \frac{\text{cov}(y_{i,g}, y_{j,g'})}{\sqrt{(\sigma^2 + \sigma_o^2)(\sigma^2 + \sigma_u^2)}}$$

$$= p(o_{i,g} = o_{j,g'}) \cdot \frac{\sigma^2}{\sigma^2 + \sigma_u^2}.$$

Weighting by population share, we have

$$\rho_w = \phi_A \rho_{w,A} + \phi_B \rho_{w,B},$$

which completes the proof. ■
B.2 Differing average incomes by caste and occupation

Lemma B.2 Assume that $\mu_{g,o}$ is allowed to vary by caste and occupation. Also assume that $\sigma_o^2 = \sigma^2, \forall o \in O$. Then,

1. $\rho_w \propto \sum_g \phi_g p(o_{i,g} = o_{j,g})$ and
2. $\rho_a \propto p(o_{i,g} = o_{j,g'})$,

both with the same constant of proportionality $\frac{\sigma^2}{\sigma_o^2 + \sigma_A^2}$.

Proof. The same argument as in the preceding lemma gives us the result. One can check that the heterogeneity in means does not affect the covariance terms. ■

B.3 Differing variances by occupation

Suppose now that

$$y_{i,g} = \mu + \epsilon_o + u_i,$$

where $\epsilon_o$ is a mean zero, variance $\sigma_o^2$ iid shock that hits each occupation. It follows that

$$\text{cov}(y_{i,A}, y_{j,B}| o_{i,A} = o, o_{j,B} = o') = 0,$$

if $o \neq o'$, and if $o = o'$ then

$$\text{cov}(y_{i,A}, y_{j,B}| o_{i,A} = o, o_{j,B} = o') = \text{E}[(\epsilon_o)(\epsilon_o)] = \sigma_o^2.$$

We still get $\text{cov}(\text{E}[y_{i,A}|o_{i,A}], \text{E}[y_{j,B}|o_{j,B}]) = 0$.

And so, we have

$$\text{cov}(y_{i,A}, y_{j,B}) = \sum_{o \in O} p(o_{i,A} = o_{j,B} = o)\sigma_o^2.$$

Thus the correlation between $y_{i,A}$ and $y_{j,B}$ is:

$$\text{corr}(y_{i,A}, y_{j,B}) = \frac{\sum_{o \in O} p(o_{i,A} = o_{j,B} = o)\sigma_o^2}{(\sum_{o \in O} p(o|A)\sigma_o)(\sum_{o' \in O} p(o'|B)\sigma_{o'})}.$$

For people in different castes, so one person from caste $A$ and one from caste $B$, we then have the following expression for across caste income correlation:

$$\rho_a = \frac{\sum_{o \in O} p(o|A)p(o|B)\sigma_o^2}{(\sum_{o \in O} p(o|A)\sigma_o)(\sum_{o' \in O} p(o'|B)\sigma_{o'})}.$$

For two people in caste $A$ we have the following expression for within caste income correlation:

$$\rho_{w,A} = \frac{\sum_{o \in O} p(o|A)^2\sigma_o^2}{(\sum_{o \in O} p(o|A)\sigma_o)^2}.$$

In sum,

$$\rho_w = p(A)\frac{\sum_{o \in O} p(o|A)^2\sigma_o^2}{(\sum_{o \in O} p(o|A)\sigma_o)^2} + p(B)\frac{\sum_{o \in O} p(o|B)^2\sigma_o^2}{(\sum_{o \in O} p(o|B)\sigma_o)^2}.$$
This implies that

\[
\rho_w - \rho_a = p(A) \frac{\sum_{o \in O} p(o|A) \sigma_o^2}{\left(\sum_{o \in O} p(o|A) \sigma_o\right)^2} + p(B) \frac{\sum_{o \in O} p(o|B) \sigma_o^2}{\left(\sum_{o \in O} p(o|B) \sigma_o\right)^2} - \frac{\sum_{o \in O} p(o|A)p(o|B)\sigma_o^2}{\left(\sum_{o \in O} p(o|A) \sigma_o\right)\left(\sum_{o' \in O} p(o'|B) \sigma_{o'}\right)}.
\]

Note that we do not know the values of \( \sigma_o \) across occupation, so we are unable to compute this measure directly from the data, as we did in previous subsections. Thus, we take two approaches. First, we proceed as before, using Measure I which ignores the variance-heterogeneity. This is an admittedly imperfect proxy for the desired measure here. Second, we treat caste as a binary random variable and occupation as a multinomial and therefore take a occupation-share weighted correlation between caste and every occupation. This is also an imperfect measure, but one that captures the intuition that if occupation can be very strongly predicted by caste, then the caste-income correlation should be higher. For this reason we use two proxies for our target quantity in the analysis.
C Supplementary Tables

Table C.1: Association between network density measures and measures of rainfall variability and within-caste income correlation

<table>
<thead>
<tr>
<th></th>
<th>Measure I</th>
<th>Measure II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Income Correlation</td>
<td>0.368 (0.149)</td>
<td>0.207 (0.148)</td>
</tr>
<tr>
<td></td>
<td>[0.006]</td>
<td>[0.24]</td>
</tr>
<tr>
<td>Income Variability</td>
<td>-0.070 (0.157)</td>
<td>-0.013 (0.157)</td>
</tr>
<tr>
<td></td>
<td>[0.668]</td>
<td>[0.974]</td>
</tr>
<tr>
<td>District FE</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Subdistrict FE</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.1384</td>
<td>0.4206</td>
</tr>
</tbody>
</table>

Notes: Outcome variable is the inverse of the average degree. Outcome variables and regressors scaled by their standard deviations. Columns (1-3) use Measure I of the caste-income correlation, whereas columns (4-6) use Measure II. Specifications include control for caste composition: \( p(1-p) \), where \( p \) is share in high caste. Wild clustered bootstrap standard errors are presented in (.), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild clustered bootstrap t are presented in [.].
Table C.2: Association between network density measures differenced across network-type and measures of rainfall variability and within-caste income correlation

<table>
<thead>
<tr>
<th>Measure</th>
<th>Measure II</th>
<th>Measure I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income Correlation x 1{Financial network}</td>
<td>0.072</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.060)</td>
</tr>
<tr>
<td></td>
<td>[0.232]</td>
<td>[0.028]</td>
</tr>
<tr>
<td>Income Variability x 1{Financial network}</td>
<td>-0.069</td>
<td>-0.060</td>
</tr>
<tr>
<td></td>
<td>(0.075)</td>
<td>(0.073)</td>
</tr>
<tr>
<td></td>
<td>[0.397]</td>
<td>[0.436]</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.9240</td>
<td>0.9270</td>
</tr>
</tbody>
</table>

Notes: Outcome variable is the inverse of the average degree of the graph of a given type (financial or social). Outcome variable, income correlation, and income variability all scaled by their standard deviations. All regressions include village fixed effects. Specifications include control for caste composition: p(1-p), where p is share in high caste, interacted with network type. Wild clustered bootstrap standard errors are presented in (), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild clustered bootstrap t are presented in [].
Table C.3: Association between average within-caste centrality of nodes with cross-caste links and measures of rainfall variability and within-caste income correlation

<table>
<thead>
<tr>
<th></th>
<th>Myerson Centrality</th>
<th>Eigenvector Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Income Correlation x 1{Financial network}</td>
<td>-0.125</td>
<td>-0.095</td>
</tr>
<tr>
<td></td>
<td>0.083</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>[0.156]</td>
<td>[0.029]</td>
</tr>
<tr>
<td>Income Variability x 1{Financial network}</td>
<td>0.290</td>
<td>-0.215</td>
</tr>
<tr>
<td></td>
<td>0.206</td>
<td>0.409</td>
</tr>
<tr>
<td></td>
<td>[0.212]</td>
<td>[0.54]</td>
</tr>
<tr>
<td>District x 1{Financial network} FE</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Subdistrict x 1{Financial network} FE</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.163</td>
<td>0.480</td>
</tr>
</tbody>
</table>

Notes: Outcome variable is the Myerson centrality, computed using our approximation algorithm, or eigenvector centrality. Income correlation given by Measure II. Outcome variables and regressors are scaled by their standard deviations. All specifications use village fixed effects and caste composition controls. Columns 2 and 4 include district-by-network type fixed effects and columns 3 and 6 include subdistrict-by-network type fixed effects. Wild clustered bootstrap standard errors are presented in (), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild clustered bootstrap t are presented in [.].