Social Investments, Informal Risk Sharing and Inequality*

Attila Ambrus[†] Arun G. Chandrasekhar[‡] Matt Elliott[§]

June 2014

Abstract

This paper investigates costly investments into social relations, 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, and examining the resulting inefficiencies of equilibrium networks. 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 also 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 society, and incomes are more correlated within groups, underinvestment is possible across groups, and more central agents have better incentives to form across group links. This reinforces the efficiency-equality trade-off. Using data from 75 Indian village networks, we provide empirical evidence consistent with predictions of our model. We verify that our effects are coming off of the financial network, by differencing out the social network.

1 Introduction

Establishing and maintaining social connections (relationships) is costly, in terms of time and other resources. However, on top of direct consumption utility, such connections can yield various economic benefits. They play a particularly highlighted role in developing countries, remedying deficiencies in market institutions. One of the most prevalent manifestations of this is that the network of social connections facilitates informal risk-sharing in the context of no formal insurance markets and limited access to lending and borrowing, a topic that has been studied extensively both theoretically and empirically.[1] Are these social investments efficient? If not, is too much time allocated to maintaining relationships or too little?

Ex ante, 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

*We thank the National Science Foundation, SES-1155302, for funding the research that collected the data we use. We thank Ben Golub and Cynthia Kinnan 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

[1]An incomplete list of papers includes Rosenzweig (1988), Fafchamps (1992), Coate and Ravallion (1993), Townsend (1994), Udry (1994), Ligon, Thomas and Worral (2000), Fafchamps and Gubert (2007), Angelucci and di Giorgi (2009).

stream which might generate improved opportunities to risk share 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 relationships 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 athropology, suggesting that members of poor communities spend inefficiently large amount of time for activities maintaining social ties, instead of productive activities. In contrast, Feigenberg et al. (2013) find evidence in a microfinence setting that it is relatively easy to experimentally intervene and create social ties among people that yield benefits to participants, suggesting the possibility of underinvestment.

To address these questions, in this paper we consider a two-stage model of network-formation and risk-sharing, in a context in which agents with CARA utilities face uncertain endowment realizations. In the first stage agents engage in a game of costly link formation, analogous to 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.[2] 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.

We assume that the social surplus generated by efficient risk-sharing arrangements is distributed among the villagers 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, at each component, then it corresponds to normative considerations: agents receive benefits proportional to their average contribution (by establishing costly links) to total surplus.[3] Second, we show that a simple decentralized procedure leads to the same outcome, providing microfoundations for the Myerson value in our setting.[4] 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 the villagers split their joint resources 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 connection have to pay higher costs towards maintaining these links, but receive no additional benefits from doing so. To address this issue we allow neighbors to engage in bilateral bargaining over

[2] Although 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 (2012). In the CARA utilities setting such unobserved income outside the scope of the risk-sharing arrangement does not affect our results.

[3] These 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).

[4] Alternative micro foundations for the Myerson value are provided by Fontenay and Gans (2004) and Navarro and Perea (2013).

state-independent transfers. For this part, we extend the canonical bargaining framework of Stole and Zwiebel (1996) to apply to any network.[5] In each pairwise negotiation a transfer is agreed that results in the two agents splitting the additional surplus generated by the link for them evenly, relative to the expected utilities they would obtain in the absence of the link. This can be thought of as the transfers being robust to renegotiation, if renegotiation would result in this 'splitting the difference' outcome. From this exercise, a recursive definition of how surpluses get divided on various networks in obtained. 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) to general networks (while simplifying their proof).

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 payoff from providing 'bridging' relationships to otherwise socially distant agents, than from providing local connections.[6] Empirical evidence supports this feature of our model - see Goyal and Vega-Redondo (2007), and references therein from the organizational literature: Burt (1994), Podolny and Baron (1997), Ahuja (2000), and Mehra et al. (2003).

In the network formation stage, we study the set of pairwise stabile networks (Jackson and Wolinsky, 1996).[7] Our first set of results concern homogenous (ex ante identical) agents. Using the inclusion-exclusion principle from combinatorics, in this case we can provide a complete characterization of stable networks. We show that for homogenous agents 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 shapes, we find that the most stable one (stable for the largest set of parameter values) is the star network, which is the most unequal one. The intuition is that that the star network minimizes the incentives of peripheral agents to establish redundant links. Conversely, the least stable efficient network is the one implying the most equal division of surplus among all stable networks.

Although agents are ex-ante identical, efficiency considerations push the structure of social relationships towards asymmetric outcomes that elevate certain individuals. Socially central individuals, who might be interpreted as community leaders, emerge endogenously from risk-sharing considerations alone.[8]

We then consider a community with multiple different groups, where within groups all agents are ex ante identical, and establishing links within groups is cheaper than across groups. We also assume that the endowment realizations of agents within groups are more positively correlated than across groups. These groups can represent different ethnic groups

[5]Stole and Zwiebel (1996) model bargaining between many employees and an employer. In terms of networks, this is a star network with the employer at the center.

[6]More 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 agent's payoff from forming a relationship are increasing in this measure.

[7]Results 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.

[8]In certain settings, over time, these central individuals could also establish more formal financial institutions, creating more entrenched inequality among agents.

or castes in a given village, or different villages.

In this more general context overinvestment remains an issue within groups, but we show that across groups underinvestment becomes possible, too, when the cost of maintaining links across groups is not too cheap. The reason is that those agents establishing the first connections across groups receive less than the total surplus generated by the connecting link, and exert a positive externality for peers in their groups. This gap between private and social benefits is smaller for agents located more centrally within their own group, providing a second force for some villagers within a group to much more central than others within the social network. For two groups, we show that the most stable efficient network shape involves star structures within groups, which are 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.

Finally, using data from 75 Indian village networks, we provide some supporting evidence for our model. We split the villagers into two groups, by caste.[9] From the theoretical analysis, risk sharing relationships are most valuable when they bridge otherwise unconnected components. And when a relationship does not provide such a bridges, its value depends on how far apart, suitable 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 the Myerson distance between any two unconnected agents, after which the pair of agents would have a profitable deviation by forming a relationship.

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 relationship between these villagers increase; The Myerson distances that can be observed in stable network therefore decrease. This results in the following predictions: (i) income variability should be positively associated with lower Myerson distances between unconnected agents; and (ii) within caste income correlation should be 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 relationship across caste. This yields our final predictions: (iii) in villages with more variable income, 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 should be lower; and (iv) in villages with more within-caste income correlation relative to across caste income correlation, the association between within-caste centrality and having an across-caste link should be lower. Because working with the exact Myerson distances is computationally infeasible, we develop an approximation of it which is exact for certain classes of graphs – trees – 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 social 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 (variability of income and within-caste income correlation) differentially vary with Myerson distance of the financial network as compared to the social

[9] There 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)).

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 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, our approach of focusing on the village-level and differencing out the social network, is extremely conservative.

On the theory side, the literature most related to our work, on social networks and informal risk-sharing agreements, consists of Bramoulle and Kranton (2007a,b), Bloch et al. (2008), Jackson et al. (2010), Billand et al. (2012) and Ambrus et al. (2014). Among these papers, Bramoulle and Kranton (2007a,b) and Billand et al. (2012) investigate network formation. The model in Bramoulle and Kranton (2007a,b) assumes that surplus on a connected income component is equally distributed, independently of the network structure. This rules out the possibility of overinvestment in their model, and leads to very different types of stable networks than in the current model. Billand et al. (2012), instead of assuming optimal risk-sharing arrangements, assume an exogenously given social norm, which prescribes that agents with a high income realization transfer a fixed amount of resources to all neighbors with low income realizations. This again leads to very different predictions on what networks should form in equilibrium.

The basic questions we investigate also arise in the social networks literature outside the risk-sharing context. For network formation models in different settings, see Jackson and Wolinsky (1996), Bala and Goyal (2000), Hojman and Szeidl (2008), and Elliott (2013). For investigations of the division of surplus in social networks, see Kranton and Minehart (2001), Calvo-Armengol (2001, 2003), Corominas-Bosch (2004), Manea (2011), and Kets et al. (2011). For a model of social investments without an explicit network structure, see Glaeser et al. (2002).

2 Preliminaries: Risk-Sharing on a Fixed Network

We consider an economy in which agents face uncertainty about the realizations of their incomes, but through the network of social connections they can redistribute their incomes, which serves as insurance from the ex ante point of view. Ultimately we are interested in investigating endogenous network formation, but in order to be able to define a noncooperative game of investing into social connections, first we will specify the risk-sharing arrangement that prevails for an exogenously given network.

The social structure

We denote the set of agents in our model by $\mathbf{N}$, and assume that they are partitioned into a set of groups $\mathbf{M}$. We let $G : \mathbf{N} \to \mathbf{M}$ be a function that assigns each person to a group; i.e., if $G(i) = g$ then person i is in group g. One interpretation of the group partitioning is that $\mathbf{N}$ represents individuals in a village, and the groups correspond to different castes. Another possible interpretation is that $\mathbf{N}$ 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 $\mathbf{N}$ such that $l_{ij} \in L$ is interpreted as a relationship exists between individuals 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 two individuals in the same group g, and between two individuals in different groups.

We will refer to the subset of $\mathbf{N}$ defined as $\mathbf{N}(i; L) = \{j : l_{ij} \in L\}$ as agent i's *neighbors*. An agent's neighbors can be partitioned according to the groups they belong to. Let $\mathbf{N}_g(i; L)$ be i's neighbors on network L from group g. We will sometimes refer subsets of agents $\mathbf{S} \subseteq \mathbf{N}$ and denote the subgraphs they generated by $L(\mathbf{S}) \equiv \{l_{ij} \in L : i, j \in \mathbf{S}\}$. A subset of agents $\mathbf{S} \subseteq \mathbf{N}$ are *path connected* on L if for each $i \in \mathbf{S}$ and each $j \in \mathbf{S}$ there exists a sequence of agents (a path) $\{i, k, k', \ldots, k'', j\}$ such that for each pair of adjacent agents in sequence (e.g. k, k'), there is a link in the network L (i.e. $l_{kk'} \in L$). For any network there is a unique partition of $\mathbf{N}$ such that there are no links between agents in different partitions but all agents within a partition are path connected. We refer to the subgraphs that makes up 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*, $d(L)$, is the maximum length of the shortest path between any two agents in that component. 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 the set $\mathbf{N}(i)$).

Incomes and consumption

Agents in $\mathbf{N}$ face uncertain income realizations. For tractability, we assume that incomes are jointly normally distributed, with expected value μ and variance σ^2 for each agent.[10] We assume that the correlation coefficient between the incomes of any two agents within the same group is ρ_w, while between incomes of any two agents not in the same group is $\rho_a < \rho_w$. That is, we assume that incomes within group are more positively correlated than across, implying that everything else being equal across group social connections have a higher potential for risk-sharing.

Although we introduce the possibility of correlated incomes in a fairly stylized way, ours is one of the first papers permitting differentially correlated incomes between different groups. Such correlations are central to the effectiveness of risk sharing arrangement, as shown below.

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

Agents can redistribute realized incomes, as described below, hence in general 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 i's consumption and λ is the coefficient of absolute risk aversion.

Efficient risk-sharing agreements

We assume that income can only be shared between individuals who are directly socially connected. However, through a sequence of bilateral transfers between connected agents, incomes can be arbitrarily redistributed within any component of the network. As the main

[10] This specification implies that we cannot impose a lower bound on the set of feasible consumption levels. As we show below, our framework generalizes readily to arbitrary income distributions, but the assumption of normally distributed shocks simplifies the analysis considerably.

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 risks efficiently, which in turn leads to ex ante Pareto-efficient risk-sharing at the level of each connected component.[11] 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.[12]

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 ω 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 for this result it is only required that agents have CARA utilities, and in particular no specific assumption is needed on the distribution of incomes.

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

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

and they are comprised of agreements of the form

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

Proof. See the Appendix. ■

The above result 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 $c_i = \frac{1}{s} \sum_{k \in \mathbf{S}} y_k + \tau_i$ is also normally distributed, and $\mathbb{E}(u(c_i)) = \mathbb{E}(c_i) - \frac{\lambda}{2} \sigma^2_{c_i}$.[13] Hence in this case the total social surplus generated by efficient

[11] 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).

[12] Kinnan (2012) 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.

[13] See for example Arrow (1965).

risk-sharing agreements is proportional to the total consumption variance reduction of the agents. This greatly simplifies computing surpluses in the analysis below.

We use $TS(L)$ to denote the total surplus generated by risk-sharing 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 one can define the incremental contribution of an agent to total surplus, as the difference between the total surplus generated by the subgraph of L defined by the given agent and those who arrived earlier, and 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 total surplus, taken over all possible orders (permutations) of players, as the player's share of the total surplus. So, letting $TS(L)$ be the expected total surplus generated on the network L, i's Myerson value is:

$$MV_i(L) \equiv \sum_{\mathbf{S} \subseteq \mathbf{N}} \frac{(|\mathbf{S}|-1)!(n-|\mathbf{S}|)!}{n!} \Big(TS(L(\mathbf{S})) - TS(L(\mathbf{S}/i)) \Big).$$

Our motivation for using the Myerson value is twofold. First, if agents on a connected component decide on 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 should be equal to the two agents that the link connects, 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 Bramoulle 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 have a pairwise meeting before the endowments realize, and negotiate a transfer. We let each agent have the option of holding up the other by deleting the link, and then negotiating its reformation so that the agents 'split the difference' and benefit equally from the relationship. Robustness to such renegotiations therefore requires that, at the marginal, each formed relationship benefits both individuals 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 $L' \subseteq L$.

Definition: For any network L, payoff vector $(u_1, ..., u_N)$ is robust to renegotiation if there is a contingent payoff scheme u given L for which:
(i) $u_i(L) = u_i$ for every $i \in N$;
(ii) $u_i(L') - u_i(L' - l_{ij}) = u_j(L') - u_j(L' - l_{ij})$ for every $i,j \in N$ and $L' \subseteq L$;
(iii) $\sum_{i \in N} u_i(L') = TS(L')$ for every $L' \subseteq L$.

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

Proposition 2 *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: $u_i(L) = MV_i(:)$.*

Proof. We will use the axiomatization of the Myerson Value established in Myerson (1980). This axiomatization states that an outcome satisfies the Myerson Value if and only if (i) each player benefits equally from a link between them; (ii) the outcome is efficient at the component level. The first property implies that on any network L:

$$MV_i(L) - MV_i(L - l_{ij}) = MV_j(L) - MV_j(L - l_{ij}).$$

The second property implies that

$$\sum_{i \in N} MV_i(L) = TS(L).$$

Myerson (1977) also establishes that the there is a unique value satisfying the above requirements, the Myerson value. Note that this system of equations is precisely the system of equations for robustness to renegotiation. ■

Our proof of Proposition 2 is a direct application of the Myerson value characterization from Myerson (1977). We therefore generalize Theorem 1 in Stole and Zweibul (96), while simplifying the proof. The Stole and Zweibul (96) result applies only to a single firm bargaining with multiple workers. In the context of the network of agreements, the structure is a star with the firm at the center. In contrast, Proposition 2 applies to any arbitrary network structure.

3 Investing Into 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. Then we conduct equilibrium analysis first in the case of only one group and then in the case of multiple groups.

3.1 The network formation game

We consider a 2-period model in which in period 1 all agents simultaneously choose which other agents they would like to form relationships with, and in the second period 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 relationship, he pays a cost $\kappa_w > 0$ if the relationship is with someone in the same group and $\kappa_a > \kappa_w$ if the relationship is with someone from a different group. This specification assumes that two agents forming a relationship 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:

$$U_i(L) = MV_i(L) - |\mathbf{N}_{G(i)}(i;L)|\kappa_w - \Big(|\mathbf{N}(i;L)| - |\mathbf{N}_{G(i)}(i;L)|\Big)\kappa_a.$$

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 $\{u_i(\mathbf{L})\}_{i\in\mathbf{N}}$ 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), 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 3 (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 $NTS(L)$:

$$NTS(L) \equiv CE\Big(\Delta \operatorname{Var}(L, \emptyset)\Big) - 2|L_w|\kappa_w - 2|L_a|\kappa_a, \tag{1}$$

where, for $L' \subset L$, $\Delta \operatorname{Var}(L, L')$ 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.[14]

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

Remark 4 *Preventing overinvestment requires ensuring that all links are essential. Any 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 – connections within a group

In this subsection 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.

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 $L(\mathbf{S}_1)$ and $L(\mathbf{S}_2)$ be the network components of agent i and agent j on network $L/\{l_{ij}\}$, and let $|\mathbf{S}_1| = s_1$ and $|\mathbf{S}_2| = s_2$. Then the sum of consumption variances on $L(\mathbf{S}_1)$ and $L(\mathbf{S}_2)$, assuming Pareto efficient risk-sharing, are $\frac{s_1+s_1(s_1-1)\rho_w}{s_1}\sigma^2$ and $\frac{s_2+s_2(s_2-1)\rho_w}{s_2}\sigma^2$. Once S_1 and S_2 are connected through l_{ij}, the sum of consumption variances on $L(\mathbf{S}_1 \cup \mathbf{S}_2)$ 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 $\Delta \operatorname{Var}(L \cup \{l_{ij}\}, L) = (1-\rho_w)\sigma^2$. 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 $\mathbf{S}_2$ when they get

[14]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).

linked to agents $\mathbf{S}_1$, as agents in the latter component can spread the income risk from agents $\mathbf{S}_2$ more effectively. On the other hand it decreases the consumption variance reduction of agents $\mathbf{S}_1$ when they get linked to agents $\mathbf{S}_2$, as the increase in s_1 implies that risk-sharing is better on $\mathbf{S}_1$ 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:

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

Since the surplus created by any essential link is $V \equiv \frac{1}{\lambda}(1 - \rho_w)\sigma^2$, 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, if 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 $\mathbf{S}$ have been added before i for a given permutation of arrival orders. As before we let $L(\mathbf{S}) \subseteq L$ be the subgraph of L such that $l_{ij} \in L(\mathbf{S})$ if and only if $l_{ij} \in L$, $i \in \mathbf{S}$ and $j \in \mathbf{S}$. If $j \notin \mathbf{S}$, then the link l_{ij} is not formed when i is added and so i receives no benefit from it. If $j \in \mathbf{S}$ 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 $L(\mathbf{S} \cup i)$ 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 $L(\mathbf{S} \cup i)$.

Remark 5 *Let $L^e(\mathbf{S} \cup i) \subseteq L(\mathbf{S} \cup i)$ be the set of essential links on $L(\mathbf{S} \cup i)$. The incremental surplus generated when i is added is proportional to the number of essential links i has on $L(\mathbf{S} \cup i)$. In other words:*

$$CE\Big(\Delta \operatorname{Var}(L(\mathbf{S} \cup i), L(\mathbf{S}))\Big) = \Big|L(i) \cap L^e(\mathbf{S} \cup i)\Big| \cdot V.$$

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 path between i and j on the network L, we let $\mathbf{P}(i,j,L) = \{P_1(i,j,L), \ldots, P_K(i,j,L)\}$ 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 $P_k(i,j,L)$.[15] We can now use these definitions to define a quantity that captures how far away two agents are on 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:

$$md(i,j,L) = \frac{1}{2} - \sum_{k=1}^{|\mathbf{P}(i,j,L)|} (-1)^{k+1} \left(\sum_{1 \leq i_1 < \cdots < i_k \leq |\mathbf{P}(i,j,L)|} \left(\frac{1}{|P_{i1} \cup \cdots \cup P_{ik}|} \right) \right)$$

[15] For example, for a path $P_k(i,j,L) = \{i, i', j\}$, $|P_k(i,j,L)| = 3$ and for paths $P_{k'}(i,j,L) = \{i, i', i'', j\}$ and $P_{k''}(i,j,L) = \{i, i', i''', j\}$, $|P_{k'}(i,j,L) \cup P_{k''}(i,j,L)| = 5$.

This expression simply 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 determines 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 and that it 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 when j is added (1/4) to the probability $P_2(i,j,L)$ is present (1/4) but then subtract the probability that both $P_1(i,j,L)$ and $P_2(i,j,L)$ are present (1/5).[16] 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 6 *If agents are ex-ante homogeneous ($m = 1$), a network L is pairwise stable if and only if*

$$\begin{aligned} &(i)\ md(i,j,L\setminus\{l_{ij}\}) \geq \kappa_w/V && \text{for all } l_{ij}\in L, \text{ and} \\ &(ii)\ md(i,j,L) \leq \kappa_w/V && \text{for all } l_{ij}\notin L. \end{aligned} \tag{2}$$

Proof. See the Appendix. ■

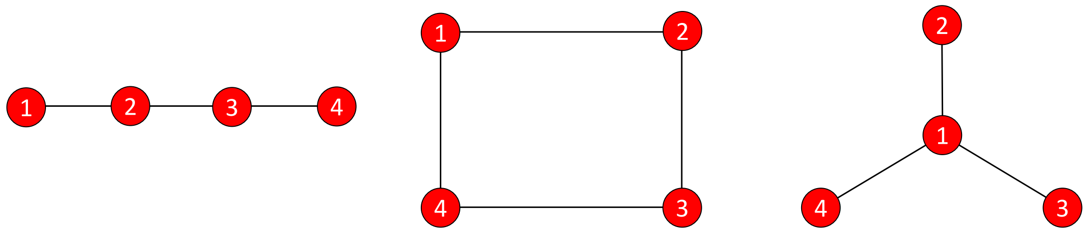
The first step in the proof of Proposition 6 establishes that the value of a link l_{ij} to i (and j) is V if the link is essential when i is added and 0 otherwise. Suppose a link l_{ij} is essential on L. It will then always induce a component reduction of one, for any orders of arrival, when the later 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 7.

Proposition 7 *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 6, 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. ■

[16] Note the both $P_1(i,j,L)$ and $P_2(i,j,L)$ are present if and only if the three nodes i', i'' and i''' are present.



(a) Efficient, unstable (b) Stable, inefficient (c) Efficient, stable

Figure 1: Stable and efficient networks for $2\kappa_w \in (\frac{2}{3}x, x)$, where x is the social value of a non-redundant link.

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

Corollary 8 *For homogeneous agents, if $2\kappa_w > V$ then the only stable network is the empty one, while if $2\kappa_w < V$ then all equilibrium networks have only one network component (all agents are path-connected).*

Corollary 9 *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.*

For the rest of the subsection we assume that $2\kappa_w < V$ so that the efficient network is not the empty network.

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 9, 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 villagers, in which there is a path from one (end) villager to the other (end) villager that passes through all other villagers exactly once (see Figure 1a). Let the star network be the unique tree network, up to a relabeling of villagers, in which one (center) villager is connected to all other villagers (see Figure 1c). On all tree networks connecting at least three villagers there are some villagers who have degree 1 (leaf nodes) and some villagers 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 depend 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 value 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 10 *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. In a tree network, by definition, there is a unique path between each pair of agents. Moreover, 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 6, 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 7 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 superflous 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 $\overline{d}(2\kappa_w) = \lceil 2V/(V - 2\kappa_w)\rceil$ complete the proof. ■

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

Corollary 11 *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

In this section we focus on the incentives for a set of path connected agents within a given group to form a social relationship outside of their own community. These different groups might correspond to people from different villages, different occupations or from different social status groups, such as castes. As, by assumption, incomes are more correlated within group than across group, there can be significant benefits from establishing such links. Moreover, we will show that these benefits accrue to all members of the community, and not just the agent establishing the across group connection. 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, there can be underinvestment, in contrast with the case of homogenous agents.

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. In, fact, it will be convenient to prove something a little more general. We will show that the link l_{ij} connects agent i from group g to agents from multiple other different groups, that $\frac{\partial \Delta \operatorname{Var}(L \cup l_{ij}, L)}{\partial s_g} > 0$. For notational convenience, set $G(i) = 0$. By definition,

$$\Delta \operatorname{Var}(L \cup l_{ij}, L) = \operatorname{Var}(L(\mathbf{S}_0)) + \operatorname{Var}(L(\mathbf{S}_1, ..., \mathbf{S}_k)) - \operatorname{Var}(L(\mathbf{S}_0, \mathbf{S}_1, ..., \mathbf{S}_k)).$$

Recalling that

$$\operatorname{Var}(L(\mathbf{S}_0, \mathbf{S}_1, ..., \mathbf{S}_k)) = \frac{\sum_{i=0}^{k}(s_i + s_i(s_i - 1)\rho_w) + 2\rho_a \sum_{i=0}^{k}(s_i \sum_{j=i+1}^{k} s_j)}{\sum_{i=0}^{k} s_i}\sigma^2,$$

it is straightforward to show that

$$\frac{\partial \Delta \operatorname{Var}(L \cup l_{ij}, L)}{\partial s_0} = 2(\rho_w - \rho_a)\frac{\sum_{i=1}^{k} s_i \sum_{j=1}^{i} s_j}{\left(\sum_{i=0}^{k} s_i\right)^2}\sigma^2 > 0,$$

which implies the claim above, since $\rho_w > \rho_a$.

An immediate implication of this observation 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. 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. This is formalized in the next result.

Let $\mathbf{S}_g = \{i : G(i) = g\}$ denote the villagers in group g.

Proposition 12 *If $m \geq 2$ then underinvestment is possible in equilibrium.*

Proof. Below we 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. First, note that if κ_w is close enough to 0, any two agents in a component have to be connected in every equilibrium. 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 $\mathbf{S}_g$ to agents in $\mathbf{S}_{g'}$ 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 than $MV(i; L \cup l_{ij}) - MV(i; L) < TS(L \cup l_{ij}) - TS(L)$ for all $i \in \mathbf{S}_g$ and $MV(j; L \cup l_{ij}) - MV(j; L) < TS(L \cup l_{ij}) - TS(L)$ for all $j \in \mathbf{S}_{g'}$. This implies that there is a range of κ_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. ■

In the general setting with many groups, there can be other types of inefficiencies in equilibrium, besides the above demonstrated underinvestment across groups. Within group links can now generate positive externalities by increasing the value of across group links and superfluous links still generate no social benefit but redistribute rents towards those establishing them. What type of inefficiencies are relevant depends on the parameter range under consideration. When κ_w is relatively low, the inefficiency of concern within group will be overinvestment and not underinvestment. A lower bound on the value of an essential link is the value generated by an essential link when no agent within the group is unconnected to any agent from another group. So, there can be no underinvestment within group whenever $\kappa_W \leq V/2$. Similarly, when κ_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. We focus our attention on this parameter region. This is also the relevant range for our data, described in the next section, where we find that within groups social connections are very dense, but across group links are considerably sparser.

Hence below we investigate what network structures within group create the best incentives to form across group links and what within group 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 against equality.

We begin by considering *local overinvestment* within group, which corresponds to the forming of superflous within group links. We found in the previous section that for homogenous agents the efficient network that minimized the incentives for overinvestment was the star. However, once we include a connection 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 is a specific such group.[17] In comparison to homogeneous agent case studied in the pervious section, the Myerson value calculation is now 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 continues to hold in a more limited sense, and the argument establishing the result is now more subtle. We focus on local network structures with only one link to other groups. Note that this is without loss of generality if there are only two groups and we are only interested in efficient networks.

Proposition 13 *Suppose a local network structure has a single across group link. The local network structure that minimizes the incentives to overinvest within group is the local star, with the across group link holder at the center. If any other local network is robust to local overinvestment, this network is also robust to local overinvestment.*

Proof. We first show that, starting from any network without an across group link, adding an across group link weakly increases the incentives for any two agents within the group to form an additional, superfluous link. To establish this, by the Myerson value calculation, we just have to show that the presence of an across group link weakly increases the variance reduction obtained by an essential within group link. This requires the following inequality:

$$\begin{aligned}\mathrm{Var}(\hat{S}_0) + \mathrm{Var}(S_0, S_1, \ldots, S_K) - \mathrm{Var}(S_0, S_1, \ldots, S_K) >\\ \mathrm{Var}(\hat{S}_0) + \mathrm{Var}(S_0) - \mathrm{Var}(\hat{S}_0, S_0) = (1-\rho_w)\sigma^2\end{aligned}$$

This holds as:

$$\begin{aligned}\mathrm{Var}(\hat{S}_0) + \mathrm{Var}(S_0, S_1, \ldots, S_K) - \mathrm{Var}(S_0, S_1, \ldots, S_K) - (1-\rho_w)\sigma^2 =\\ \frac{\hat{s}_0\left(\sum_{k=2}^{K} s_k^2 + \sum_{k=2}^{K-1} s_k \sum_{j=k+1}^{K} s_j\right) 2(\rho_w - \rho_a)\sigma^2}{\sum_{k=1}^{K} s_k \left(\hat{s}_0 + \sum_{k=1}^{K} s_k\right)} > 0.\end{aligned}$$

So the incentives to overinvest are weakly larger once the across group link is added. The overinvestment incentives without the across group link provide a lower bound on the possible overinvestment incentives with the across group link. We will show that the local star, with the across group link holder at the center, achieves this lower bound. The presence of the across group link does *not* increase the incentives for overinvestment within group. As by

[17] This contrasts with the case of an essential across group link that bridges two otherwise disconnected groups. In this case, we showed that the variance reduction is increasing in the size of the groups connected, allowing us to conclude that there can be underinvestment in such links.

Corollary 11 the local star minimized overinvestment incentives absent the across group link, it must then also minimize overinvestment incentives with the across group link. In other words, once the across group link is added (to the center node) the incentives to overinvest are no higher for this network, but are weakly higher for all other efficient networks, and so the star must still minimize the incentives for overinvestment.

We now complete the proof by showing that the payoffs of two periphery nodes i, j in the local star increase by the same amount when they form a (superfluous) link, regardless of whether of whether the center node has an across group link or not. Suppose this additional link is formed. 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 to the network. 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 either 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. ■

We now consider the local network structures that maximize the incentives for an across group link to be established. In the proof of Proposition 12 we showed that the marginal contribution to total surplus from a 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 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 $\mathbf{P}(\mathbf{S}_k)$ be the set of possible permutations for agents $\mathbf{S}_k$. For any permutation $P \in \mathbf{P}(\mathbf{S})$, let $\mathbf{T}_i(P)$ be the set of agents i is path connected to on $L(\mathbf{S}')$ where $\mathbf{S}'$ is the set of agents including i drawn weakly before i. Let $T_i^{(m)}$ be a random variable that equals the cardinality of $\mathbf{T}_i(P)$, conditional on individual i being the m-th person to be drawn according to P.

We will say a person $i \in \mathbf{S}_k$ is more central within group than $j \in \mathbf{S}_k$, if $T_i^{(m)}$ first-order stochastically dominates $T_j^{(m)}$ for all $m \in \{1, 2, ..., |\mathbf{S}_k|\}$. In other words, considering all the permutations such that i and j is the m-th agent to be drawn, 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.[18] This measure of centrality provides a partial ordering of agents.

Proposition 14 *Suppose agents in $\mathbf{S}_0$ form a network component. Suppose each other villager j is from some other group ($j \notin \mathbf{S}_0$) and these other villagers form a network component. Let $i, i' \in \mathbf{S}_0$ and let $j \notin \mathbf{S}_0$. If i is more central within group than i', then i receives a higher payoff from forming the link l_{ij} than i' receives from forming the link $l_{i'j}$:*

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

Proof. See the Appendix. ■

The proof of Proposition 12 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

[18] An alternative equivalent definition is that i is more central than j if there exists a bijection $B : \mathbf{P}(\mathbf{S}_k) \to \mathbf{P}(\mathbf{S}_k)$ such that $|\mathbf{T}_i(P)| \geq |\mathbf{T}_j(B(P))|$ and $P(i) = P'(j)$, where $P(i)$ is the position of i in the permutation P and $P' = B(P)$.

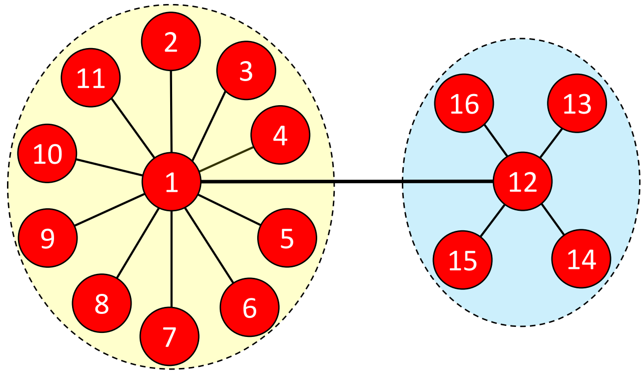


Figure 2: Agents 1 to 11 are in one group and agents 12 to 16 in the other. The local network structure within each group is a star with agents 1 and 12 at the respective centers. These agents also have an across group link.

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 12 shows that more central agents have better incentives to form a bridging link that will connect their group to other groups. We can then consider the problem of maximizing the incentives for links to be formed across groups by choosing the local network structures within groups. We will say that the local network structures (networks containing only within group links) that achieve these maximum possible incentives are most robust to underinvestment inefficiency across group.

Corollary 15 *The efficient local network structure most robust to underinvestment inefficiency is the local star, with the potential across group link holder at the center. If any other local network is robust to underinvestment across group, this network is also robust to underinvestment across group.*

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 that all other agents and so by Proposition 12 has stronger incentives to form an across group link.

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

The above results establish that for two groups, the efficient networks that minimize overinvestment within group and underinvestment accross group are center-connected stars, as in Figure 2.

The above results further reinforce the tension between efficiency and equality. The local star not only minimizes the incentives for overinvestment within group, it also minimizes the incentives for underinvestment across group. 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.[19] 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 for the other agents.[20]

6 Empirical analysis

We apply our model to study risk-sharing networks in Indian villages. 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. Murdoch (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.

Our model is ideal for describing risk-sharing behavior within and across castes in a village. In the language of our model, every caste corresponds to a different group and it is ex ante more costly for individuals to form cross-caste links. This is in line with an extensive sociological literature (see e.g., Srinivas (1962)). The model we have developed provides direction as to how cross-caste insurance networks should vary as a function of economic environmental factors we have data pertaining to, such as variability of income or the within-versus-across caste income correlation.

To investigate these issues, we make use of a unique dataset of detailed network data from 75 villages in Karnataka, India. The dataset is particularly well-suited for our analysis as (i) it involves numerous independent villages (essential for statistical inference), (ii) it includes *complete* network data across both financial and social relationships across all households (in fact every adult individual) in every village (network-based studies are notoriously subject to measurement error), and (iii) caste is salient in these communities.

6.1 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. The survey included a village questionnaire, a census of all households, demographic covariates (including caste, sub-caste, and occupation), as well as data on a number of amenities (such as roofing, latrine or electricity access quality). A detailed individual level survey was administered to all adults in every village and the survey included a social networks module with twelve dimensions of relationships including financial relationships, social relationships, and advice relationships.[21]

[19] By Corollary 15, the local star maximizes the 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 by assumption $V > 2\kappa_w$, and we saw in the proof of Proposition 14 that the value of an essential within group links increases when there is an across group link.

[20] The central agent still receives half the increase in value of the within group links. Inequality will increase if and only if the central agent would be willing to form the across group link even if there were no within group links.

[21] See Banerjee et al. (2014) for more details.

Our analysis focuses on two types of networks in the analysis: the *financial graph*, L^F, and the *social graph*, L^S. The financial graphs represent risk-sharing relationships. From the data we build an "AND" network, which says a link exists if it exists on every dimension being considered (various types of financial relationships).[22] The advantage of doing this is that it generates a network structure that is more robust to independent measurement error. Turning to the social graph, for symmetry we again focus on an "AND" network, saying that a link exists if nodes have relationships across every social dimension.[23] 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 is only about 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 the core fissure in the South Indian social fabric.

We study how the network structure of L^F varies with parameters of the economic environment. We are chiefly interested in income variance, σ^2, and the within-caste relative to across-caste income correlation $(\rho_w - \rho_a)$. To track income variability, we use a variable corresponding to the variability of rainfall in the village. We do this by merging 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 rainfally by village. As described below, we show that our results are robust to defining rainfall variability as the village-based standard deviation or by looking at the standard deviation once removing month fixed-effects.

Measuring income correlation is difficult. Ideally, we would have time series data on incomes of all households allowing us to calculate the income correlation both within and across caste. 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 within the same occupation within a community will be correlated. We 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 villagers in the occupation. Formally, for a given village we consider:

$$\sum_{k=1}^{K} \operatorname{corr}\left(Caste, Occupation_k\right) \cdot \mathrm{P}\left(Occupation_k\right).$$

This constructs a measure which is 0 if there is no correlation between caste group and occupation and 1 if caste group perfectly predicts occupational choices. We argue that higher values of this within versus across caste occupation correlation is a good proxy for the within versus across caste income correlation, in the absence of other data.[24]

[22] 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.

[23] 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.

[24] As income correlations within a given occupation will likely differ by village, separate estimates of within

6.2 Empirical Framework

6.2.1 Predictions

In order to see whether the our data is consistent with our model, we investigate several distinct predictions from the theory.

Proposition 6 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. Specifically, for every pair of villagers i and j that do not have a link:

$$md(i, j, L) \leq \frac{\kappa_w \lambda}{(1-\rho_w)\sigma^2}. \tag{3}$$

We note that Proposition 6 has been constructed for the case where the network consists of one group. In our empirical setting, we are interested in Indian village networks where the clear abstraction is that there are multiple groups, given by caste.

Nevertheless, inequality 3 provides an appropriate benchmark for within group links. Recall that the left hand side of inequality 3 captures the probability that the link is not essential for a random permutation in the arrival order and the right hand 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 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.

Letting $s_1, \ldots, s_K$ be the distribution of people from different groups in the first component, and $s'_1, \ldots, s'_K$ be the distribution of people from different groups in the second component, the value of the variance reduction is:[25]

$$\frac{\kappa_w \lambda}{(1 - f(s_1, \ldots, s_k, s'_1, \ldots, s'_k)\rho_w - g(s_1, \ldots, s_k, s'_1, \ldots, s'_k)\rho_a)\sigma^2}.$$

We already know that when $s_i = s'_j = 0$ for all $i, j \neq 1$, $f(s_1, \ldots, s_k, s'_1, \ldots, s'_k) = 1$ and $g(s_1, \ldots, s_k, s'_1, \ldots, s'_k) = 0$. However, the same simplification occurs when the distribution of people from each group is the same within each component, so $s_i = \alpha$ for $i = 1, \ldots, k$ and $s_j = \beta$ for $j = 1, \ldots, k$.

Proposition 16 *The certain equivalent value of variance reduction obtained by connected a component with α people from groups 1 to k with a component with β people from groups 1 to k is:*

$$\frac{\kappa_w \lambda}{(1-\rho_w)\sigma^2}.$$

Proof. See the Appendix. ■

Proposition 16 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 relative number of people from each group is the same

and across caste income correlations would be noisier. By looking at within-versus-across caste income correlation we difference out some of these level effects and to get a less noisy get a more accurate measure.

[25] This can be observed by manipulating the equations in the proof of Proposition 12.

in each component. Moreover, the across group income correlation does not matter. While asymmetries can increase or reduce this variance reduction, equation 3 provides a useful approximation for an upper bound on agents' Myerson distances.

Recalling that we can use our data to estimate $\rho_w - \rho_a$ better than we can estimate, ρ_w alone, the above considerations lead to the following predictions:

P1. In villages with higher σ^2, the Myerson distance between two villagers should be smaller on average.

P2. In villages with lower $\rho_w - \rho_a$, the Myerson distance between two villagers should be smaller on average.

While our first set of predictions looks at the relationship of the Myerson distance and the economic environmental parameters, our second set of predictions looks at the composition of the links. Our interest here is which agents provide the across group links. Proposition 14 shows that more central agents have better incentives to provide an across group link. How important centrality will be will depend on the overall strength of incentives to form across caste links. When income variance is very high, or within caste income correlation is very high relative to across caste income correlation, the incentives to form across caste link will be very 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 equations in the proof of Proposition 12 it is straightforward to show that for an across caste bridging link l_{ij}:

$$\frac{\partial \Delta \operatorname{Var}(L, L \cup \{l_{ij}\})}{\partial \sigma^2} > 0 \qquad \frac{\partial \Delta \operatorname{Var}(L, L \cup \{l_{ij}\})}{\partial \rho_w} > 0 \qquad \frac{\partial \Delta \operatorname{Var}(L, L \cup \{l_{ij}\})}{\partial \rho_a} < 0.$$

This means that the incentives to form an across caste link are increasing in σ^2 and $\rho_w - \rho_a$, leading to the following predictions:[26]

P3. In villages with higher σ^2, the association between within caste centrality and providing across caste links is stronger.

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

6.2.2 Approximating the Myerson distance

Our next task is to compute the Myerson distance of every pair in every village. As it turns out, this is computationally infeasible for the sample sizes of our data.

Let $\mathbf{md}(L)$ be the matrix of Myerson distances and define $\mathbf{q}(L) = 1/2 - \mathbf{md}(L)$. So $\mathbf{q}(L)$ is a matrix with the ijth entry capturing the probability that upon being added to the network agent i will not be connected to agent j. It can be thought of as a measure of density of L. Finally, let $\mathbf{A}$ be the adjacency matrix corresponding to L.

It is difficult to directly characterize $\mathbf{md}(L)$ (or equivalently, $\mathbf{q}(L)$). Given that each village typically consists of around 230 households, there are 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 6), and it has to be

[26] Recall that with our data we can estimate $\rho_w - \rho_a$, but not ρ_w and ρ_a separately.

done for all unconnected pairs of agents i and j. In addition, due to presumed measurement error (see discussion in Banerjee et al. (2013)), there are likely to be missing paths. In fact, the data have occassional disconnected components and therefore measures that are precisely based on exact paths or even maximal path lengths are likely to be problematic (Chandrasekhar and Lewis (2014)).

Instead, we develop a computationally feasible approximation of $\mathbf{md}(L)$. Our approximation is exactly $\mathbf{md}(L)$ for tree graphs. For robustness, we also make use of two standard network measures to approximate the sparseness of a network as characterized by $\mathbf{md}(L)$: inverses of both the average degree, $\bar{D}(L) := n^{-1} \sum_j \sum_i A_{ij}$ and the maximal eigenvalue $\lambda_1(\mathbf{A})$ of the graph.

To approximate $\mathbf{q}$, we use the following idea. The inclusion-exclusion principal weights a path of length l by $(1/2)^{l-1}$. It then combines these 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 up and substract minimal paths that share some nodes in common. This second step is complicated and we will abstract from it. However, we will be able to include only minimal paths and to also weight these paths appropriately. We denote our approximation of $\mathbf{q}$ by $\hat{\mathbf{q}}$

Algorithm 17 (Approximation of q) *Let e^i be the ith basis vector. Initialize $\widehat{\mathbf{q}}$ to be a $n \times n$ matrix of zeros. Let $z^{t,i}$ and $x^{t,i}$ index n-vectors initialized as zeros as well for $i = 1, ..., n$ and $t = 1, ..., n$.*

1. *Period 1:*
   (a) *Percolation: $x^{1,i} = \mathbf{A} e^i$.*
2. *Period 2, given $(x^{1,i}, \mathbf{A})$:*
   (a) *Zeroing columns: $z^{2,i} = e^i$.*
   (b) *Update matrix: $\mathbf{A}_2 = zeros(n),\ \mathbf{A}_2(:, \sim z^{2,i}) = \mathbf{A}(:, \sim z^{2,i})$.*
   (c) *Percolation: $x^{2,i} = \mathbf{A}_2 x^{1,i}$.*
3. *Period t, given $(x^{t-1,i}, A_{t-1})$:*
   (a) *Zeroing columns: $z^{t,i} = \mathbf{1}\left\{\sum_{s=2}^{t} x^{s-2,i} > 0\right\}$.*
   (b) *Update matrix: $\mathbf{A}_t = zeros(n),\ \mathbf{A}_t(:, \sim z^{t,i}) = \mathbf{A}_{t-1}(:, \sim z^{t,i})$.*
   (c) *Percolation: $x^{t,i} = \mathbf{A}_t x^{t-1,i}$.*
4. *End if $x^{t,i} = x^{t-1,i}$. Let the final period be T.*
5. *Set $\widehat{q}_{ij} = \sum_{t=1}^{T} x_j^{t,i} \left(\frac{1}{t+1}\right)$ where $x_j^{1,i}$ is the jth entry of $x^{1,i}$.*[27]
6. *Repeat steps (1-5) for all $(e^1, ..., e^n)$ to compute $\widehat{\mathbf{q}}$.*

This measure is exact for trees.

Proposition 18 *Let L be a connected tree. Then $\widehat{\mathbf{q}}(L) = \mathbf{q}(L)$.*

[27] Set $\widehat{q}_{ij} = 0$ if $l_{ij} \in L$.

Proof. See the Appendix. ■

The approximation $\widehat{\mathbf{q}}(L)$ works by treating all minimal paths as independent while some rely on the same nodes. This means that in general it weakly overestimates q_{ij}.

To operationalize this in our village level regression analysis, we use $\widetilde{q}(L) := \sum_{i<j} \widehat{q}_{ij}/\binom{n}{2}$ which measures an appropriately weighted density of the network.

6.2.3 Empirical approach

Our predictions are about how network structure varies with parameters σ^2 or $\rho_w - \rho_a$. Most of our analysis is observational (not causal) and simply looks at the cross-sectional variation of network structure with these parameters through OLS. By focusing on different aspects of network structure (either the Myerson distance or the composition of cross-caste links) and how it varies with the economic environmental parameters, we hope to provide a wide array of evidence in support of our predictions.

One key feature we exploit is the fact that our theory is built for risk-sharing networks and not, say, for social relationships. This allows us to take a stringent differences-in-difference approach and can study whether the correlations we document come from L^F, the financial graph, and as opposed to L^S, the social graph. That is, we view the theory as being *differentially more informative* about the structure of financial links as opposed to, say, social links and we view the difference in patterns across link-types to be informative.[28] Furthermore, as unobserved heterogeneity 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 may allow us to test our theory with more power. As our predictions are specifically about risk-sharing relationships, we will be able to exploit the multigraph by examining whether our predictions are *differentially* at play in the financial networks relative to the social networks.

We note that our empirical approach is more conservative than similar studies in the literature (i.e., Karlan et al. (2007), Ambrus et al. (2014), Kinnan and Townsend (2014)) in terms of statistical inference. These studies typically have very few networks (i.e., using 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. Additionally, they 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 focusing on village-level variation, we are allowing for arbitrary correlation within graphs and, 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.

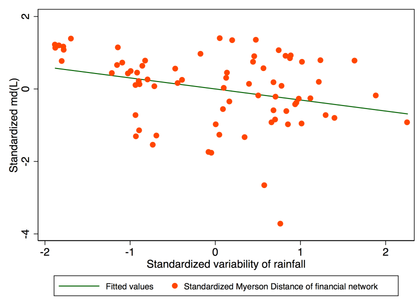
6.3 Results

6.3.1 Myerson distance as a function of income variability and income correlation

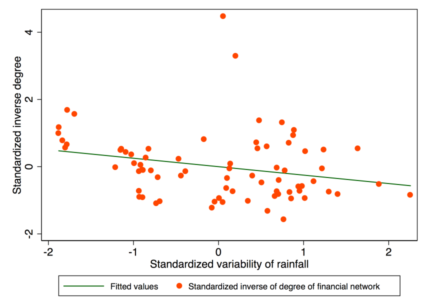
We begin with the predictions P1 and P2. In villages with higher σ^2, is the average Myerson distance between two villagers smaller? In villages with lower ρ_w, is the average Myerson distance between two villagers smaller?

[28] Of course, different link-types may be viewed in the context of a larger dynamic game.

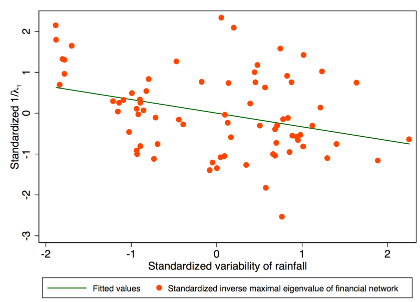
Figure 3 presents our results for income variance graphically. As predicted, we see that villages in areas corresponding to more variable income processes are associated with lower Myerson distance measured by $\widetilde{md}(L^F)$. This is robust to the two alternative metrics of network sparsity we use: $1/\bar{D}(L^F)$ and $1/\lambda_1(\mathbf{A}^F)$. This supports the theory as with more variable incomes longer distances cannot be supported in a stable network.



(a) Measure: $\widetilde{md}(L^F)$



(b) Measure: $1/\bar{D}(\mathbf{A}^F)$



(c) Measure: $1/\lambda_1(\mathbf{A}^F)$

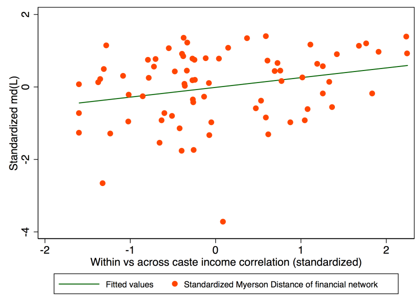
Figure 3: Measures of network sparsity in standardized units against the variability of rainfall at the village level in standardized units.

Next we turn to a similar analysis for within versus across caste income correlation. Figure 4 presents our results graphically. As predicted, we see that villages with relatively more within caste income correlation is associated with higher $\widetilde{md}(L^F)$. Again, this result is robust to looking at the inverse of both average degree and maximal eigenvalue.

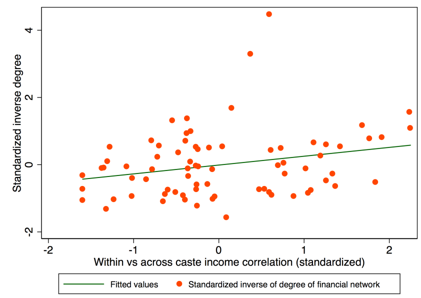
Table 1 demonstrates the robustness of this graphical evidence in a simple regression analysis. Here we show that the results are mostly robust to a multivariate regression – conditional on both σ and $\rho_w - \rho_a$ – as well as with subdistrict fixed effects and controls for caste composition.

Columns 1-3 present the raw conditional correlations. We see that a one standard deviation increase in rainfall variability is associated with a 0.3 standard deviation decrease in $\widetilde{md}(L^F)$ (column 1). Similarly a one standard deviation increase in the caste income correlation measure is associated with a 0.26 standard deviation increase in $\widetilde{md}(L^F)$. We see similar results turning to the other measures of network sparsity: $1/\bar{D}(L^F)$ and $1/\lambda_1(\mathbf{A}^F)$ (columns 2-3).

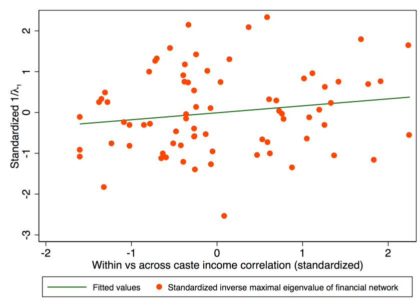
We then look within taluk (subdistrict) by including taluk fixed effects and also include



(a) Measure: $\widetilde{md}(L^F)$



(b) Measure: $1/\bar{D}(\mathbf{A}^F)$



(c) Measure: $1/\lambda_1(\mathbf{A}^F)$

Figure 4: Measures of network density in standardized units against the within versus across caste income correlation metric in standardized units.

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

	(1) md(L)	(2) 1/D(L)	(3) 1/λ(L)	(4) md(L)	(5) md(L)	(6) 1/D(L)	(7) 1/D(L)	(8) 1/λ(L)	(9) 1/λ(L)
Rainfall variability	-0.299***	-0.245***	-0.331***	-0.337***	-0.329**	-0.115	-0.270***	-0.158	-0.295*
	(0.0913)	(0.0763)	(0.0964)	(0.0967)	(0.144)	(0.188)	(0.0776)	(0.184)	(0.158)
Within vs across income correlation	0.262***	0.257***	0.163*	0.132**	0.126**	0.171	0.173	-0.0328	-0.0322
	(0.0858)	(0.0777)	(0.0867)	(0.0590)	(0.0604)	(0.124)	(0.122)	(0.121)	(0.116)
Taluk FE	N	N	N	Y	Y	Y	Y	Y	Y
Caste Composition Control	N	N	N	Y	Y	Y	Y	Y	Y
Observations	75	75	75	75	75	75	75	75	75
R-squared	0.158	0.126	0.137	0.742	0.740	0.683	0.687	0.712	0.717

Notes: All outcome variables and regressors scaled by their standard deviations. Columns (1-3), (4), (6) and (8) use the standard deviation of rainfall by village. Columns (5), (7), (9) use the standard deviation of rainfall (removing month fixed effects). Cluster bootstrapped standard errors, by taluk, with 1000 samples are used.

control for the caste composition (columns 4-9).[29] We also present results varied by whether we use the standard deviation of rainfall by village (columns 4, 6, 8) or we use the standard deviation of rainfall where we have removed month fixed effects by village and show that our results are robust (columns 5, 7, 9). Our results for $\widetilde{md}(L^F)$ remain robust (columns 4, 5): a

[29]This is simply $\sqrt{p(1-p)}$, the standard deviation of the caste distribution where p denotes the probability of being SC/ST.

standard deviation increase in rainfall variability is associated with a 0.33 standard deviation drop in the average Myerson distance and a standard deviation increase in caste income correlation is associated with a 0.13 standard deviation drop in the Myerson distance. The results for inverse of degree and the maximal eigenvalue are more mixed – either reflecting the prediction or indicating that the effect is too noisy to be distinguished from zero.

Taken together, we have provided evidence that increases in income variability and within-caste income correlation are associated with network density measures in a manner consistent with our theory.

6.3.2 Differences by network type

Next we exploit the fact that the theory we developed pertains to risk-sharing network. Using a difference-in-differences approach, we check whether the effects we are interested are coming differentially from L^F, the financial graph, as opposed to the social graph, L^S. This is a very demanding test because there are likely to be fixed-cost reasons as to why having a social link generates a financial link for free (or more cheaply) or vice-versa.

To do this graphically, we generate a density measure that is the difference between the financial graph and the social graph and look at the relationship with income variability or income correlation variables.

Figure 5 presents the results graphically. We see that villages in areas corresponding to more correlated within-caste income processes are associated with greater $\widetilde{md}(L^F) - \widetilde{md}(L^S)$. 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.

To operationalize this in a regression, letting v index village and $t \in \{F, S\}$ index network type, we look at

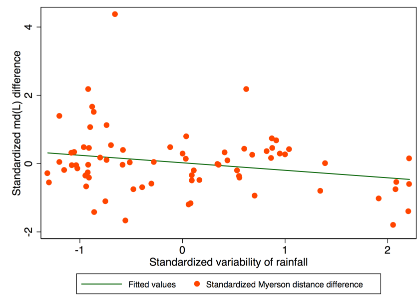
$$Y_{v,t} = \alpha + \beta \cdot \sigma_v \cdot \mathbf{1}_{\{t=F\}} + \gamma \cdot (\rho_{w,v} - \rho_{a,v}) \cdot \mathbf{1}_{\{t=F\}} + \mu_v + \delta \cdot X_v \cdot \mathbf{1}_{\{t=F\}} + \epsilon_{v,t}.$$

The crucial prediction under our model is that $\beta < 0$ 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. This also allows for taking an arbitrary village-level fixed effect μ_v as well as type-village level controls.

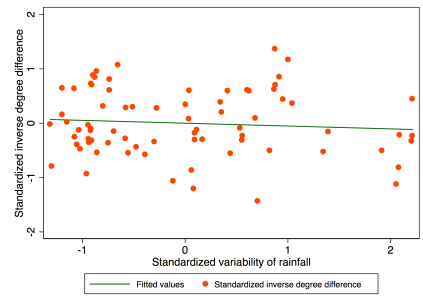
Table 2 presents the results. All specifications include village fixed effects, so we are making comparisons about the relative shape of the financial versus social network for a fixed village as we vary the economic parameter of interest. We find that a one standard deviation increase in rainfall variability differentially decreases $\widetilde{md}(L)$ by 0.2 standard deviations more in financial networks than social networks (column 1). Similarly, a one standard deviation increase in caste income correlation differentially increases $\widetilde{md}(L)$ by 0.23 standard deviations more in financial networks than social networks (column 1). These results are robust to inclusion of a caste composition by network type control.

We repeat the exercises for both the inverse of the average degree and the maximal eigenvalue. We robustly find that caste income correlation is differentially more positively associated with the measure of network sparsity for the financial networks, irrespective of specification. We are unable to reject no association between income variability and these alternative measures of network sparsity, differentially across network type.

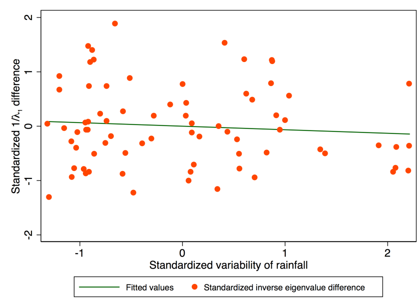
The evidence presented here illustrates that even when we look within-village, by allowing for village fixed effects, we see that the financial graphs behave in a manner consistent with the theory in a way that the social graphs do not.



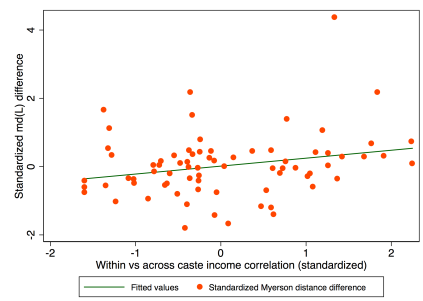
(a) $\widetilde{md}(L^F) - \widetilde{md}(L^S)$



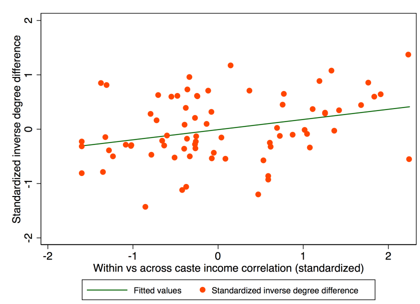
(b) $1/\bar{D}(\mathbf{A}^F) - 1/\bar{D}(\mathbf{A}^S)$



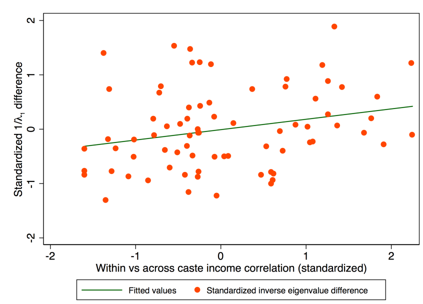
(c) $1/\lambda_1(\mathbf{A}^F) - 1/\lambda_1(\mathbf{A}^S)$



(d) $\widetilde{md}(L^F) - \widetilde{md}(L^S)$



(e) $1/\bar{D}(\mathbf{A}^F) - 1/\bar{D}(\mathbf{A}^S)$



(f) $1/\lambda_1(\mathbf{A}^F) - 1/\lambda_1(\mathbf{A}^S)$

Figure 5: Measures of density, differenced across network type, versus income variability or within-caste income correlation.

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

The results until now have looked at how the average Myerson distance is associated with economic environmental variables. We now look at how the composition of cross caste links vary with these parameters. We are interested in how the centrality of those holding across caste links varies with income correlation within caste and the variance of rainfall. P3 demonstrates that in villages with higher variability we should see a greater association between within-caste centrality and having an across caste link. P4 shows that a similar effect is true for looking at within versus across caste income correlation.

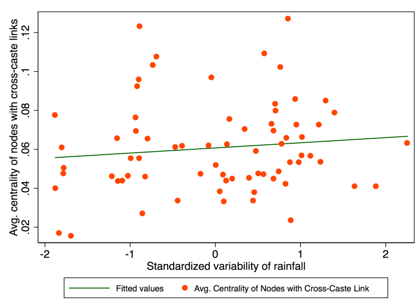
Figure 6 presents the results graphically. We observe that once we difference the social graph, we only see a negative relationship between within-caste income correlation and the

Table 2: Association between network density measures differenced across network-type and measures of rainfall variability and within-caste income correlation

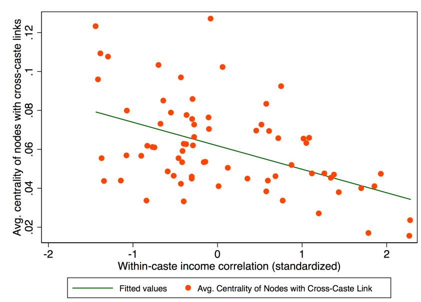
	(1) md(L)	(2) 1/D(L)	(3) 1/λ(L)	(4) md(L)	(5) 1/D(L)	(6) 1/λ(L)
Rainfall variability x **1**{Financial network}	-0.214**	-0.0468	-0.0635	-0.216**	-0.0480	-0.0647
	(0.0926)	(0.0598)	(0.0772)	(0.0953)	(0.0630)	(0.0744)
Income correlation x **1**{Financial network}	0.226*	0.187***	0.193**	0.216**	0.180***	0.186**
	(0.116)	(0.0660)	(0.0835)	(0.108)	(0.0677)	(0.0844)
Village FE	Y	Y	Y	Y	Y	Y
Caste Composition x **1**{Financial network}	N	N	N	Y	Y	Y
Observations	150	150	150	150	150	150
R-squared	0.809	0.919	0.872	0.811	0.919	0.873

Notes: Outcome variable is stated measure of network sparsity for a given network type (financial network or social network). Variability of rainfall is computed as the standard deviation of rainfall in a village having removed month dummies. Standard errors are bootstrap clustered at the village level with 1000 repitions.

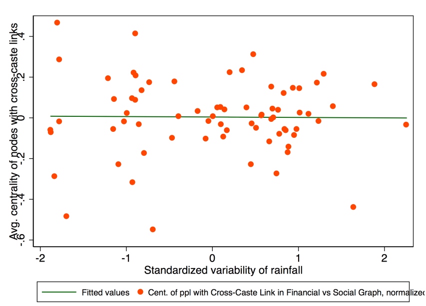
average centrality of those with cross-caste links.



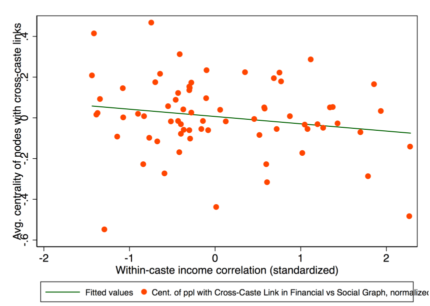
(a) Variability of income



(b) Within-caste income correlation



(c) Variability of income differencing financial vs social graph



(d) Within-caste income correlation differencing financial vs social graph

Figure 6: How the association between centrality within caste and propensity to link across caste varies with environment.

To investigate this formally we look at a regression of the average eigenvector centrality of the set of nodes in a community that have cross-caste links against income variability and within-caste correlation. Table 3 presents the results. We look at regressions without fixed effects and controls for caste composition (columns 1, 3) as well as with both (columns 2, 4). Additionally, we look at effects in levels – where the outcome variable is the average centrality of nodes with cross-caste links in the financial network – as well as in differences – where we difference the outcome variable across the financial and social graphs. Here we use our differences-in-differences approach, where we ask whether the effect differentially persists for financial graphs. Furthermore, our specification, with taluk fixed effects when regressing outcome variable of the difference in centralities can be interpreted as allowing for taluk-type fixed effects, which is even more stringent than our initial difference-in-difference specification.

We robustly and significantly find that an increased within-caste income correlation is associated with a decline in the average centrality of nodes with cross-caste links. A one standard deviation increase in the within-caste income correlation is associated with between a 0.34 to a 0.52 standard deviation decline in the centrality of the nodes in cross-caste links.

Table 3: Association between average within-caste centrality of nodes with cross-caste links and measures of rainfall variability and within-caste income correlation

	(1)	(2)	(3)	(4)
	Average centrality of cross-caste linkers		Average centrality of cross-caste linkers (Financial - Social)	
Rainfall variability	0.1005	0.1333	0.0802	-0.3265
	(0.1121)	(0.2893)	(0.1095)	(0.2246)
Within vs across income correlation	-0.5196	-0.4862	-0.4389	-0.3412
	(0.1401)	(0.1401)	(0.1050565	(0.1122)
Taluk FE	N	Y	N	Y
Caste Composition Control	N	Y	N	Y

Notes: For ease of interpretation, outcome variables as well as variability of rainfall and income correlation variables are scaled by their standard deviations. Outcome variable in columns (1-2) is the average eigenvector centrality of cross-caste linkers computed from the financial graph. The outcome variable in columns (3-4) is the difference of the eigenvector centrality of cross-caste linkers computed from the financial and social graphs.

We are unable to either confirm or reject our theory when we look at rainfall variability. We are unable to reject that there is no association, and the most promising evidence only comes forth when we look at the difference of the financial graph and the social graph and look within the subdistrict (column 4, $p = 0.174$).

Thus, our evidence is partially in support of our theory. We find that indeed the association between increasing the income correlation and decreasing the relationship between centrality and cross-caste linking is borne out in the data under this test. However, we are neither able to confirm nor reject our hypothesis when looking at the variability of income.

6.4 Discussion

In this section, we have provided suggestive evidence, consistent with our theory. We have demonstrated the following:

1. Networks exhibit lower average Myerson distance (or other measures of sparsity) when there is more variable rainfall.

2. Networks are higher average Myerson distance (or other measures of sparsity) when there is more within versus across group income correlation.

3. Networks exhibit stronger relationships 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 argue the effect is coming from a graph that measures risk-sharing behavior (the financial graph) as opposed to an alternative type of network relationship.

We highlight that our results are strongest when we use a network measure motivated directly by our theory, $\widetilde{md}(L)$, as compared to standard measures of network sparsity.

Having several distinct predictions such as this is useful as we are conducting an observational analysis. The fact that our predictions are generally being supported off the financial network rather than social network, also helps gives us confidence that the relevant explanation is about risk sharing. There are several other stories consistent with the first prediction; for instance, any theory where the returns to investing in risk-sharing relationships go up should predict more dense networks. We do note, however, that this prediction does not hold in Bramoulle and Kranton (2007a), which is perhaps the closest paper to ours. Furthermore, our compositional story that differentiates between risk sharing within and across caste and permits higher within than across caste income correlation is new. The second, third and forth predictions are either specific 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 caveat that these statements are clearly observational (non-causal) and subject to measurement error. However, we note that 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. Previous work on risk-sharing networks typically have a single or a handful of networks and rarely exploit the multigraph data. Additionally, we are extremely conservative in conducting statistical inference, by only relying on the independence across-villages and not exploiting any of the within-village observations (which are likely to be correlated). The previous literature, on the other hand, treats dyads or nodes as independent or – at least – has to assume extremely limited within-network correlation for their statistical inference.

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. We find that overinvestment into social relations is likely in a homogeneous community, but there is potential underinvestment into social connections bridging agents in different groups when incomes are correlated less positively across than within groups. We find a novel trade-off between equality and efficiency. Within groups the social structure that has the highest level of inequality is the star. However, 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 superflous within group links that redistribute but do not create surplus.

Using a unique dataset of 75 Indian villages, we also find empirical support for our model. 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 as in 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 including a dynamic extension of the analysis that allows for autocorrelation between income realizations as well.

References

Ambrus, A., M. Mobius, and A. Szeidl (2014): "Consumption risk-sharing in social networks," American Economic Review, 104, 149-182.

Angelucci, M. and G. De Giorgi (2009): "Indirect effects of an aid program: How do cash injections affect non-eligibles' consumption?," American Economic Review, 99, 486–508.

Arrow, K. (1965): "Aspects of the Theory of Risk-Bearing (Yrjo Jahnsson Lectures)," Yrjo Jahnssonin Saatio, Helsinki.

Austen-Smith, D. and R. Fryer (2005): "The economic analysis of acting white," Quarterly Journal of Economics, 120, 551-583.

Bala, V. and S. Goyal (2000): "A noncooperative model of network formation," Econometrica, 68, 1181-1229.

Banerjee, A., A. Chandrasekhar, E. Duflo and M. Jackson (2013): "The diffusion of microfinance," Science, 341, No. 6144 1236498

Banerjee, A., A. Chandrasekhar, E. Duflo and M. Jackson (2014): "Gossip: Identifying central individuals in a social network," MIT and Stanford working paper.

Bergstrom, T. and H. Varian (1985): "When do markets have transferable utility?," Journal of Economic Theory, 35, 222-233.

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

Borch, K. (1962): "Equilibrium in a reinsurance market," Econometrica, 30, 424-444.

Bramoulle, Y. and R. Kranton (2007a): "Risk-sharing networks," Journal of Economic Behavior and Organization, 64, 275-294.

Bramoullé, Y. and R. Kranton (2007b): "Risk sharing across communities," American Economic Review Papers & Proceedings, 97, 70.

Calvo-Armengol, A. (2001): "Bargaining power in communication networks," Mathematical Social Sciences, 41, 69-87.

Calvo-Armengol, A. (2002): "On bargaining partner selection when communication is restricted," International Journal of Game Theory, 30, 503-515.

Calvo-Armengol, A. and R. Ilkilic (2009): "Pairwise-stability and Nash equilibria in network formation," International Journal of Game Theory, 38, 51-79.

Cole, H. L., and N. R. Kocherlakota (2001): "Efficient allocations with hidden income and hidden storage," The Review of Economic Studies 68, 523-542.

Coate, S. and M. Ravaillon (1993): "Reciprocity without commitment: Characterization and performance of informal insurance arrangements," Journal of Development Economics, 40, 1–24.

Corominas-Bosch, M. (2004): "Bargaining in a network of buyers and sellers," Journal of Economic Theory, 115, 35-77.

Elliott, M. (2013): "Inefficiencies in networked markets," mimeo CalTech.

Fafchamps, M. (1992): "Solidarity networks in preindustrial societies: Rational peasants with a moral economy," Economic Development and Cultural Change, 41, 147-74.

Fafchamps, M. and F. Gubert (2007): "The formation of risk sharing networks," Journal of Development Economics, 83, 326-350.

Feigenberg, B., E. Field and R. Pande (2013): "The economic returns to social interaction: Experimental evidence from micro finance," Review of Economic Studies, 80, 1459-1483.

Fontenay, C. and J. Gans (2013): "Bilateral bargaining with externalities," mimeo University of Toronto.

Glaeser, E., D. Laibson, and B. Sacerdote (2002): "An economic approach to social capital," The Economic Journal, 112, F437-F458.

Hojman, D. and A. Szeidl (2008): "Core and periphery in networks," Journal of Economic Theory, 139, 295-309.

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

Jackson, M. (2010): "Social and economic networks," Princeton University Press.

Jackson, M. and A. Wolinsky (1996): "A strategic model of social and economic networks," Journal of Economic Theory, 71, 44-74.

Kets, W., G. Iyengar, R. Sethi and S. Bowles (2011): "Inequality and network structure," Games and Economic Behavior, 73, 215-226.

Matsuura, K. and C.J. Willmott (2012): "Terrestrial precipitation: 19002008 gridded monthly time series,"Center for Climatic Research Department of Geography Center for Climatic Research, University of Delaware.

Mazzocco, Maurizio (2012): "Testing efficient risk sharing with heterogeneous risk preferences,"The American Economic Review, 102.1, 428-468.

Munshi, K. and Rosenzweig, M. (2006): "Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy,"The American Economic Review, 1225-1252.

Munshi, K. and Rosenzweig, M. (2009): "Why is mobility in India so low? Social insurance, inequality, and growth,"National Bureau of Economic Research, (No. w14850).

Kinnan, C. and R. Townsend (2012): "Kinship and financial networks, formal financial access, and risk reduction," The American Economic Review, 102, 289-293.

Kranton, R. and D. Minehart (2001): "A Theory of buyer-seller networks," American Economic Review, 91, 485-508.

Ligon, E., , J. Thomas, and T. Worrall (2002): "Informal insurance arrangements with limited commitment: Theory and evidence from village economies," Review of Economic Studies, 69, 209–244.

Manea, M. (2011): "Bargaining in stationary networks," American Economic Review, 101, 2042-80.

Myerson, R. (1977): “Graphs and cooperation in games,” Mathematics of Operations Research, 2, 225-229.

Navarro, N. and A. Perea, (2013): “A simple bargaining procedure for the Myerson value,” The BE Journal of Theoretical Economics, 13, 1-20.

Rosenzweig, M. (1988): “Risk, implicit contracts and the family in rural areas of low-income countries,”Economic Journal, 98, 1148-70.

Stole, L. and J. Zwiebel (1996): “Intra-firm bargaining under non-binding contracts,”Review of Economic Studies, 63, 375-410.

Townsend, R. (1994): “Risk and insurance in village India,” Econometrica, 62, 539-591.

Udry, C. (1994): “Risk and insurance in a rural credit market: An empirical investigation in Northern Nigeria,” Review of Economic Studies, 61, 495-526.

Wilson, R. (1968): “The theory of syndicates,” Econometrica, 36, 119-132.

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:

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

This shows that the amount K' needed to compensate i from keeping the more stochastic consumption stream c_i instead of consumption steam 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 $\omega \in \Omega$:

$$\frac{\frac{\partial u_i(c_i(\omega))}{\partial c_i(\omega)}}{\frac{\partial u_j(c_j(\omega))}{\partial c_j(\omega)}} = \alpha_{ij}$$

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

$$\begin{aligned}
\frac{e^{-\lambda c_i(\omega)}}{e^{-\lambda c_j(\omega)}} &= \alpha_{ij} \\
c_i(\omega) - c_j(\omega) &= -\frac{\ln(\alpha_{ij})}{\lambda} \\
\mathbf{E}[c_i(\omega)] - \mathbf{E}[c_j(\omega)] &= -\frac{\ln(\alpha_{ij})}{\lambda} \\
c_i(\omega) - c_j(\omega) &= \mathbf{E}[c_i(\omega)] - \mathbf{E}[c_j(\omega)] \qquad (4)
\end{aligned}$$

Letting i and j be neighbors, such that $j \in \mathbf{N}(i)$, equation 4 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 $\mathbf{S}$ in each state of the world to minimize the sum of their consumption variances:

$$\min_{\mathbf{c}} \sum_{i \in \mathbf{S}} Var(c_i)$$

subject to $\sum_{i \in \mathbf{S}} y_i(\omega) = \sum_{i \in \mathbf{S}} c_i(\omega)$ in all possible states of the world ω. Note,

$$\sum_{i\in\mathbf{S}}\text{Var}(c_i) = \sum_{i\in\mathbf{S}}\sum_{\omega\in\Omega} p(\omega)(c_i(\omega) - \mathbf{E}[c_i])^2$$

where $p(\omega)$ is the probability of state ω. 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 $i \in \mathbf{S}$ and each $\omega \in \Omega$:

$$2(c_i(\omega) - \mathbf{E}[c_i]) = \gamma(\omega),$$

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

$$c_i(\omega) - c_j(\omega) = \mathbf{E}[c_i(\omega)] - \mathbf{E}[c_j(\omega)]$$

for all $i, j \in \mathbf{S}$. 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:

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

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

Proof of Proposition 6. 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 $\widehat{\mathbf{S}}_i \subseteq \mathbf{N}$ identifies the set of agents, including i, who are drawn weakly before i in the permutation. We let $\mathcal{L}(i,L)$ be the probability distribution over the networks $L(\widehat{\mathbf{S}}_j)$, 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 $L(\widehat{\mathbf{S}})$ drawn from $\mathcal{L}(i,L)$.

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 Myserson 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

$$q(i,j,L) = \sum_{k=1}^{|\mathbf{P}(i,j,L)|}(-1)^{k+1}\left(\sum_{1\le i_1<\cdots<i_k\le|\mathbf{P}(i,j,L)|}\left(\frac{1}{|P_{i1}\cup\cdots\cup P_{ik}|}\right)\right),$$

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 $\mathbf{P}(i,j,L') \neq \emptyset$; 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 $P_k(i,j,L) \in \mathbf{P}(i,j,L)$ 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 $\mathcal{L}(j,L)$. 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 $\mathcal{L}(j,L)$, i.e., $\cup_{P_k(i,j,L)\in\mathbf{P}(i,j,L)} Pr(P_k(i,j,L))$. As these paths are not disjoint, the inclusion-exclusion principle needs to be applied. Doing so results in the formula shown.[30] ■

Proof of Proposition 16. Recall that

$$\mathrm{Var}(s_1,\ldots,s_k) = \sigma^2 \left(\frac{\sum_{i=1}^{k}(s_i + s_i(s_i-1)\rho_w) + 2\rho_a \sum_{i=1}^{k}(s_i \sum_{j=i+1}^{k} s_j)}{\sum_{i=1}^{k} s_i} \right). \tag{5}$$

We need to calculate $\mathrm{Var}(\alpha,\ldots,\alpha) + \mathrm{Var}(\beta,\ldots,\beta) - \mathrm{Var}(\alpha+\beta,\ldots,\alpha+\beta)$. Applying the above formula yields:

$$\begin{aligned}
&\sigma^2 \left(\frac{k(\alpha + \alpha(\alpha-1)\rho_w) + 2\rho_a\alpha^2((k-1)+(k-2)+\cdots+1)}{k\alpha} \right) \\
&+ \sigma^2 \left(\frac{k(\beta + \beta(\beta-1)\rho_w) + 2\rho_a\beta^2((k-1)+(k-2)+\cdots+1)}{k\beta} \right) \\
&- \sigma^2 \left(\frac{k(\alpha+\beta + (\alpha+\beta)(\alpha+\beta-1)\rho_w) + 2\rho_a(\alpha+\beta)^2((k-1)+(k-2)+\cdots+1)}{k(\alpha+\beta)} \right).
\end{aligned}$$

This simplifies to:

$$\begin{gathered}
((1+(\alpha-1)\rho_w) + 2\rho_a\alpha((k-1)/2))\,\sigma^2 \\
+ ((1+(\beta-1)\rho_w) + 2\rho_a\beta((k-1)/2))\,\sigma^2 \\
- ((1+(\alpha+\beta-1)\rho_w) + 2\rho_a(\alpha+\beta)((k-1)/2))\,\sigma^2 = \\
(1-\rho_w)\sigma^2,
\end{gathered} \tag{6}$$

which completes the proof. ■

Proof of Proposition 18. 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.

[30] 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: $(1/K) + (1/K') - (1/K'')$.

Consider the implementation of the algorithm to find $\widehat{q}_{ij}$. We begin by calculating $x^{1,i} = \mathbf{A}e^i$, where e^i be the ith basis vector. This identifies all agents connected to i. We then delete the ith column from the adjacency matrix $\mathbf{A}$ and set this new matrix equal to $\mathbf{A}_2$. This deletes the outward links from i in the network L. Starting from i's neighbors we then find their neighbors on $\mathbf{A}_2$. In other words we calculate $x^{2,i} = \mathbf{A}_2 x^{1,i}$. This identifies the distance 2 neighbors of i. We then delete the columns of $\mathbf{A}_2$ that are indexed by one of i's neighbors and so on. In the tth round the algorithm identifies the distance t neighbors of i. Thus, for all $t < l$, $x_j^{l,i} = 0$, for $t = l$, $x_j^{l,i} = 1$ and for all $t > l$, $x_j^{t,i} = 0$. Thus, if the unique path from i to j is length l, $\widehat{q}_{ij} = 1/l$. From equation [x] it is also easily checked that $q_{ij} = 1/l$. ■

Proof of Proposition 14. Consider the set of arrival order permutations for all agents for the network $L \cup l_{i\prime j}$. 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 $i\prime$ in the original permutation; (ii) when added i connects to exactly the same set of agents $\mathbf{N} \setminus \mathbf{S}_0$ as $i\prime$ connects to; (iii) i is connected to weakly more agents within $\mathbf{S}_0$ when added than $i\prime$. In the proof of Proposition [x – underinvestment possible], it was shown that the risk reduction, and hence the additional payoff to $k \in \mathbf{S}_0$, from the across-group link l_{kj} is an increasing function of the component size of k's groups. It then follows that:

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

To match permutations we do the following. Take the set of original permutations for $L \cup l_{i\prime j}$ and make the following adjustments. First, switch the arrival positions of $i\prime$ and i. This alone is enough to ensure that condition (i) and (ii) are satisfied. Now note that as i is more central than $i\prime$ the CDF over agents i is connected to within $\mathbf{S}_0$ upon being added first order stochastically dominates the set of agents that $i\prime$ is connected to within $\mathbf{S}_0$. Label the permutations for $i\prime'$ of agents in $\mathbf{S}_0$ in the following way. First consider all permutations in which $i\prime$ is added first. Label these permutations $P_{i\prime}(1) \ldots$. Now consider all permutations in which one other agent is added before $i\prime$ and continue to label permutations by ordering these permutation in terms in ascending order of how many agents $i\prime$ 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 $\mathbf{S}_0$ for $i\prime$ is permutation $P_{i\prime}(x)$, reorder the arrival of agents within $\mathbf{S}_0$ to be $P_i(x)$. This permutation for i now also satisfies condition (iii). By construction, i will be connected to more agents within $\mathbf{S}_0$ than $i\prime$ 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. ■