NBER WORKING PAPER SERIES

SOCIAL INVESTMENTS, INFORMAL RISK SHARING, AND INEQUALITY

Attila Ambrus
Arun G. Chandrasekhar
Matt Elliott

Working Paper 20669
http://www.nber.org/papers/w20669

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2014

We thank the National Science Foundation, SES-1155302, for funding the research that collected the data we use. Ambrus, Chandrasekhar and Elliott acknowledge financial support from National Science Foundation (NSF) grant "The economic benefits of investing into social relationships," SES-1429959. We thank Nageeb Ali, Ben Golub, Matt Jackson, Willemien Kets, Cynthia Kinnan, Rachel Kranton, Peter Landry, Horacio Larreguy, Melanie Morten, Kaivan Munshi and Luigi Pistaferri for helpful comments. We also thank seminar participants at the Calvo-Armengol prize conference, the Harvard-MIT theory seminar, Oxford, Microsoft Research, 2012 ThRed conference in Budapest, Syracuse, Ohio State University, Duke and the Tinbergen Institute. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Social Investments, Informal Risk Sharing, and Inequality
Attila Ambrus, Arun G. Chandrasekhar, and Matt Elliott
NBER Working Paper No. 20669
November 2014
JEL No. C78,D31,D61,D86,L14,Z13

ABSTRACT

This paper studies costly network formation in the context of risk sharing. Neighboring agents negotiate agreements as in Stole and Zwiebel (1996), which results in the social surplus being allocated according to the Myerson value. We uncover two types of inefficiency: overinvestment in social relationships within group (e.g., caste, ethnicity), but underinvestment across group. We find a novel tradeoff between efficiency and equality. Both within and across groups, inefficiencies are minimized by increasing social inequality, which results in financial inequality and increasing the centrality of the most central agents. Evidence from 75 Indian village networks is congruent with our model.

Attila Ambrus
Department of Economics
Duke University
213 Social Sciences Building
419 Chapel Drive, Campus Box 90097
Durham, NC 27708
aa231@duke.edu

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

Matt Elliott
Division of the Humanities and Social Sciences 228
California Institute of Technology
Pasadena, California 91125
melliott@hss.caltech.edu

1. Introduction

In the context of missing formal insurance markets and limited access to lending and borrowing, incomes may be smoothed through informal risk-sharing arrangements that utilize social connections. A large theoretical and empirical literature studies this topic,[1] but considerably less attention has been paid to the social investments that enable risk sharing, how risk-sharing networks form in the presence of these costly investments, and the resulting implications for efficiency and equity.[2] At the same time, there is growing empirical evidence that risk-sharing networks respond to financial incentives,[3] and that in general risk-sharing networks form endogenously, in a way that depends on the economic environment. A central question we address is whether such investments are efficient, and if not then whether too many or too few resources are allocated to maintaining relationships. We are also interested in whether the resulting equilibrium forces generate social inequality (asymmetries in network positions) in society, even when agents are ex-ante homogeneous, and if social inequality translates into financial inequality.

Both underinvestment and overinvestment in social capital are conceivable. Two people establishing a social connection to share risk gain access to a less stochastic income stream which might generate improved opportunities to share risk with their other connections. As these positive spillovers might not be fully taken into account when deciding whether to establish the link, underinvestment can prevail. On the other hand, if more socially connected individuals receive a higher share of the surplus generated by risk sharing, that can lead to overinvestment. Villagers may form links to redistribute the surplus towards themselves, rather than to increase the overall surplus generated. The empirical literature also suggests that both types of inefficiencies are possible, in different contexts. Austen-Smith and Fryer (2005) cites numerous references from sociology and anthropology, suggesting that members of poor communities allocate inefficiently large amounts of time to activities maintaining social ties, instead of productive activities. In contrast, Feigenberg et al. (2013) find evidence in a microfinance setting that it is relatively easy to experimentally intervene and create social ties among people that yield substantial benefits. One explanation for this finding is that there is underinvestment in social relationships.

It is important to study whether there is too little or too much investment into social relations, both to put related academic work (which often takes social connections to be

[1]An incomplete list of papers includes Rosenzweig (1988), Fafchamps (1992), Coate and Ravallion (1993), Townsend (1994), Udry (1994), Ligon, Thomas and Worral (2002), Fafchamps and Gubert (2007), Bloch, Genicot, and Ray (2008), Angelucci and di Giorgi (2009), Jackson, Rodriguez-Barraquer and Tan (2012), Ambrus, Mobius and Szeidl (2014).

[2]Previous works that does consider the network formation problem include Bramoulle and Kranton (2007a,b) in the theoretical literature and Attanasio et al. (2012) in the experimental literature. For a related paper outside the networks framework, see Glaeser et al. (2002).

[3]See recent work by Binzel et al. (2014) and Banerjee et al. (2014b,c), which in different contexts look at how social networks respond to the introduction of financial instruments such as savings vehicles or microfinance.

exogenously given) into context and to guide policy choices. Consider the example of microfinance. If there is overinvestment, microfinance has a greater scope for efficiency savings in terms of reducing people's allocation of time into social investments. With underinvestment, however, it has more scope for smoothing incomes. If there is neither under- nor overinvestment, it also tells us that informal risk sharing is working relatively well as a second-best solution. Understanding which regime applies can help anticipate policy implications and evaluate welfare impacts of interventions.

To explore efficiency and inequality, in this paper we consider a two-stage model of network formation and risk sharing, in a context in which agents with constant absolute risk aversion (CARA) utilities face uncertain endowment realizations. In the first stage, agents choose with whom to form connections. Link formation is costly, as in Myerson (1991) and Jackson and Wolinsky (1996). In the second stage, connected agents commit to a risk-sharing arrangement that is contingent on future endowment realizations.[4] We show that in our CARA setting, expected utilities are transferable through state-independent transfers—and, up to these transfers, efficient risk-sharing arrangements on any network component are uniquely pinned down. The latter determine the allocation of surplus among agents. To keep the model tractable, we abstract away from the issues of how to enforce risk-sharing agreements.[5]

We assume that the social surplus generated by efficient risk-sharing arrangements is distributed among the agents according to the Myerson value, a network-specific version of the Shapley value.[6] Our motivation here comes from two sources. First, if the surplus division is chosen in a centralized manner, then it has normative appeal on the grounds of fairness: two agents benefit equally from a social relationship between them, and receive benefits proportional to their average contributions to total surplus (from establishing costly links).[7] Second, we show that a simple and natural decentralized procedure leads to the same outcome, providing microfoundations for the Myerson value in our setting. The procedure combines the exchange algorithm of Bramoulle and Kranton (2007a) and the pairwise robustness to renegotiation requirement of Stole and Zwiebel (1996). Following Bramoulle and Kranton (2007a), efficient risk sharing is obtained by assuming that after every realization of endowments, there is an infinite sequence of pairwise exchanges between neighbors in which joint resources are split equally. This process equalizes the consumptions of the agents on a connected component of the network. However, given this social norm, all connected agents receive the same

[4]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 (2011). In the CARA utilities setting, such unobserved income outside the scope of the risk-sharing arrangement does not affect our results.

[5]See for example Ambrus et al. (2014) for an investigation of such issues.

[6]For investigations of the division of surplus in social networks in other contexts, see Calvo-Armengol (2001, 2003), Corominas-Bosch (2004), Manea (2011), and Kets et al. (2011).

[7]These motivations make the Myerson value a commonly used concept in the network formation literature. See a related discussion on pp. 422–425 of Jackson (2010).

consumption independent of the structure of the social network. This means, for example, that agents with more social connections have to pay higher costs towards maintaining these links but receive no additional benefits from doing so. To address this issue, we allow neighbors to engage in bilateral bargaining over state-independent transfers. These agreements are made ex ante, that is prior to income realizations. For this part, we extend the canonical bargaining framework of Stole and Zwiebel (1996) to arbitrary networks.[8] In each pairwise negotiation, the two linked agents agree to a transfer that evenly splits the surplus generated by their link over and above the expected payoffs they would receive in it's absence. This can be thought of as the transfers being robust to renegotiation if renegotiation would result in a "splitting the difference" outcome. From this exercise we obtain a recursive definition of how surpluses get divided on different networks. By applying the axiomatization of the Myerson value provided in Myerson (1980), we show that the unique division of surplus compatible with this recursive definition is the Myerson value. In doing so, we provide new foundations for the Myerson value.[9] Moreover, we do so simply by applying the axiomatization provided by Myerson (1980).

A key implication of the Myerson value determining the division of surplus is that more centrally connected agents receive a higher share of the surplus. Moreover, in our risk-sharing context it implies that agents receive larger payoffs from providing "bridging links" to otherwise socially distant agents than from providing local connections.[10] Empirical evidence supports this feature of our model—see Goyal and Vega-Redondo (2007), and references therein from the organizational literature: Burt (1992), Podolny and Baron (1997), Ahuja (2000), and Mehra et al. (2001).

In the network formation stage, we study the set of pairwise-stable networks (Jackson and Wolinsky, 1996).[11]

Our general analysis considers a community comprised of different groups where all agents within each group 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. Groups can represent different ethnic groups or castes in a given village, or in different villages. The core results show that there can be

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

[9]The process is decentralized, as it envisages pairwise renegotiations. The result in Stole and Zwiebel (1996) that we extend is Theorem 1. Their Theorem 2 can also be extended to our setting, and this would provide fully noncooperative foundations. Indeed, related noncooperative foundations are provided by Fontenay and Gans (2013), while Navarro and Perea (2013) take a different approach to microfounding the Myerson value. Slikker (2007) also provides noncooperative foundations, although the game analyzed is not decentralized: offers are made at the coalitional level.

[10]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 agents' payoffs from forming a relationship are increasing in this measure.

[11]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.

overinvestment within groups but not underinvestment, whereas across groups underinvestment is likely to be the main concern.

To see the intuition about overinvestment within groups, we first consider the case of homogeneous agents, that is, when there is only one group. Using the inclusion–exclusion principle from combinatorics,[12] we provide a complete characterization of stable networks. We show that in this case there can never be underinvestment in social connections, as agents establishing an essential link (connecting two otherwise unconnected components of the network) always receive a benefit exactly equal to the social value of the link. However, overinvestment, in the form of redundant links, is possible, and becomes widespread as the cost of link formation decreases. We also find a trade-off between efficiency and equality. Among all possible efficient network structures, we find that the most stable (in the sense of being stable for the largest set of parameter values) is the star, which also results in the most unequal division of surplus. The intuition is that the star network minimizes the incentives of peripheral agents to establish redundant links. Conversely, the least stable efficient network entails the most equal division of surplus among all stable networks. Although agents are ex-ante identical, efficiency considerations push the structure of social connections towards asymmetric outcomes that elevate certain individuals. Socially central individuals emerge endogenously from risk-sharing considerations alone.[13]

Turning attention to the case of multiple groups, we find that across-group underinvestment becomes an issue when the cost of maintaining links across groups is sufficiently high.[14] The reason is that the agents who establish the first connection across groups receive less than the total surplus generated by the link, providing positive externalities for peers in their groups. This gap between private and social benefits is smaller for agents located more centrally in their own group, providing a second force for some agents within a group to be more central. For two groups, we show that the most stable efficient network structure involves stars within groups, connected by their centers, and we establish a weaker form of this result for more than two groups. This reinforces the trade-off between efficiency and equality in the many-groups context.

Using data from 75 Indian villages, we provide some supporting evidence for our model. We split the villagers into two groups, by caste.[15] From the theoretical analysis, risk-sharing links are most valuable when they bridge otherwise unconnected components. And when a link does not provide such a bridge, its value depends on how far apart, suitably defined, the agents would otherwise be on the social network. We call this distance between a pair of agents their Myerson distance. Our theory predicts that there is an upper bound on the

[12]See Chapter 10 in van Lint and Wilson (2001).

[13]In certain settings, these central individuals might establish more formal financial institutions over time, creating more entrenched inequality.

[14]While across-group overinvestment remains possible, the main concern when across-group link costs are relatively high is underinvestment.

[15]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)).

Myerson distance between any two unconnected agents within the same group, beyond which the pair of agents would have a profitable deviation by forming a link. In addition, we predict that there will be inquality in social positions and that more central agents within their group will form across-group links. However, there are many alternative stories consistent with these predictions. We therefore also generate more subtle predictions to consider: how changes in the economic environment in terms of income variability or correlation correspond to changes in network structure under our model, while differencing out the friendship network from the risk-sharing network.

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

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

Ultimately, our empirical approach allows us to be more conservative than similar studies in the literature (e.g., Karlan et al. (2009), Ambrus et al. (2014), and Kinnan and Townsend (2012)). While the studies above have access to just a few of networks (e.g., 2, 1 and 16, respectively), we have 75 networks and also multigraph data. Most studies, therefore, are forced to do statistical inference within networks, which limits the number degree of correlated shocks they are able to handle. Relative to this approach, our focus on the village level and differencing out the social network is extremely conservative.

On the theory side, the studies on social networks and informal risk sharing that are most related to ours include Bramoulle and Kranton (2007a,b), Bloch et al. (2008), Jackson et al. (2012), Billand et al. (2012), Ali and Miller (2013a,b), and Ambrus et al. (2014). Many of these papers focus on the enforcement issues we abstract from and investigate how social capital can be used to sustain cooperation for lower discount factors than would otherwise be possible. We take a complementary approach and instead focus on the distribution of surplus and the incentives this creates for social investments. One way of viewing our approach to enforcement is that we instead consider the part of the parameter range where the discount factor is sufficiently high to sustain cooperation without levering social capital. Among the aforementioned papers, Bramoulle and Kranton (2007a,b) and Billand et al. (2012) investigate costly network formation. Bramoulle and Kranton's (2007a,b) model assumes that the surplus on a connected income component is equally distributed, independently of the network structure. This rules out the possibility of overinvestment, and leads to different types of stable networks than in our model. Instead of assuming optimal risk-sharing arrangements, Billand et al. (2012) assume an exogenously given social norm, which prescribes that high-income agents transfer a fixed amount of resources to all low-income neighbors. This again leads to very different predictions regarding the types of networks that form in equilibrium.

More generally, network formation problems are important. Establishing and maintaining social connections (relationships) is costly, in terms of time and other resources. However, on top of direct consumption utility, such links can yield many economic benefits. Papers studying formation in different contexts include Jackson and Wolinsky (1996), Bala and Goyal (2000), Kranton and Minehart (2001), Hojman and Szeidl (2008), and Elliott (2013). Notable in this literature is a lack of empirical work, which can be attributed to a number of innate difficulties that taking these models to data presents. One common problem is a multiplicity of stable networks. But perhaps most important is that the networks in question can only very rarely be partitioned into a sizeable number of separate networks that can reasonably be treated as independent. Since predictions are often at the level of the overall network structure, this makes testing extremely challenging. Our many observations of social networks that are relatively independent of each other, coupled with our approach to circumventing data limitations, allow us to provide a first step towards testing predictions based on the overall network structure. And although we study a quite specific network formation problem tailored to risk sharing in villages, the general structure of our problem is relevant to other applications.[16]

[16]For a different and more specific application, suppose researchers can collaborate on a project. Each researcher brings something heterogenous and positive to the value of the collaboration, so that the value of the collaboration is increasing in the set of agents involved. Collaboration is possible only when it takes place

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

2. Preliminaries: Risk Sharing on a Fixed Network

To study social investments and the network formation problem we need to first specify how formed networks affect payoffs. Taking this backward induction approach, we begin by considering risk sharing on a fixed network. We consider an economy in which agents face stochastic income realizations but can insure against this uncertainty through redistributions made over the network of social connections. Ultimately, we are interested in investigating endogenous network formation. However, in order to define a noncooperative game of investing into social connections, first we will specify the risk-sharing arrangement that prevails for any given network.

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 agent to a group; i.e., if $G(i) = g$ then agent i is in group g. One interpretation of the group partitioning is that $\mathbf{N}$ 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 the existence of a link between agents i and j. The social network influences both the set of feasible risk-sharing arrangements and the distribution of surplus from risk sharing, as described below. There can be links both between agents in the same group and between agents in different groups.

We will refer to $\mathbf{N}(i; L) \equiv \{j : l_{ij} \in L\} \subset \mathbf{N}$ 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. A *path* is a sequence of agents $\{i, k, k', \dots, k'', j\}$ such that every pair of adjacent agents in the sequence is linked. The *path length* of a path is the number of agents in the path. We will sometimes refer to subsets of agents $\mathbf{S} \subseteq \mathbf{N}$ and denote the subgraphs they generate by $L(\mathbf{S}) \equiv \{l_{ij} \in L : i, j \in \mathbf{S}\}$. A subset of agents $\mathbf{S} \subseteq \mathbf{N}$ is *path connected* on L if, for each $i \in \mathbf{S}$ and each $j \in \mathbf{S}$, there exists a path connecting i and j. 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 cells of this partition as *network components*. A *shortest path* between two path-connected agents

among agents who are directly connected to another collaborator and surplus is split according to the Myerson value (as in our work, motivated by robustness to renegotiations). Such a setting fits into our framework.

i and j is a path connecting i and j with a lower path length than any other. The *diameter of a network component*, $d(C)$, is the maximum value—taken over all pairs of agents in C—of the length of a shortest path. 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 $\mathbf{N}(i; L)$).

Incomes and Consumption

Agents in $\mathbf{N}$ face uncertain incomes realizations. For tractability, we assume that incomes are jointly normally distributed, with expected value μ and variance σ^2 for each agent.[17] We assume that the correlation coefficient between the incomes of any two agents within the same group it is ρ_w, while between incomes of any two agents not in the same group is $\rho_a < \rho_w$.[18] That is, we assume that incomes are more positively correlated within groups than across groups, so that all else equal, social connections across groups have a higher potential for risk sharing.

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

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

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

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

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

Efficient Risk-Sharing Agreements

We assume that income cannot be directly shared between agents $i, j \in \mathbf{N}$ unless they are connected, i.e., $l_{ij} \in L$. However, through a sequence of bilateral transfers between connected agents, incomes can be arbitrarily redistributed within any component of the network. As the main focus of our paper is network formation, to keep the model tractable we abstract away from enforcement constraints, and analogously to Bramoulle and Kranton (2007a, b), we assume that all neighboring agents share risk efficiently, which in turn leads to ex-ante Pareto-efficient risk sharing at the level of each connected component. While in practice risk sharing is imperfect, perfect 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 the

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

[18]It is well-known that for a vector of random variables, not all combinations of correlations are possible. We implicitly assume that our parameters are such that the resulting correlation matrix is positive semidefinite.

remaining income is privately observed and never shared. Results are very similar for this more general setting.[19]

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

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

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

and they are comprised of agreements of the form

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

where $\tau_i \in \Re$ is a state independent transfer.

Proof. See Appendix A. □

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

Proposition 1 implies that the total surplus generated by an efficient risk-sharing arrangement is an increasing function of the reduction in the sum of consumption variances. For a general distribution of shocks, this function can be complicated. However, if shocks are jointly normally distributed, then $c_i = \frac{1}{|\mathbf{S}|} \sum_{k \in \mathbf{S}} y_k + \tau_i$ is also normally distributed, and

[19]Kinnan (2011) finds evidence that hidden income can explain imperfect risk sharing in Thai villages relative to the enforceability and moral hazard problems we are abstracting from. Cole and Kocherlakota (2001) show that when individuals can privately store income, state-contingent transfers are not possible and risk sharing is limited to borrowing and lending.

$\mathrm{E}(u(c_i)) = \mathrm{E}(c_i) - \frac{\lambda}{2}\mathrm{Var}(c_i)$.[20] Hence in this case the total social surplus generated by efficient risk-sharing agreements is proportional to the total consumption variance reduction. This greatly simplifies the computation of surpluses in the analysis below.

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

Division of Surplus

We assume that agents on a connected component divide the total surplus created by the risk-sharing arrangement according to the Myerson value (Myerson (1977), (1980)). The Myerson value is a cooperative solution concept defined in transferable utility environments that is a network-specific version of the Shapley value. The basic idea behind it is the same as for the Shapley value. For any order of arrivals of the players, the incremental contribution of an agent to the total surplus can be derived as the difference between the total surpluses generated by the subgraph of L defined by the given agent and those who arrived earlier, and by the subgraph that is defined by only those agents who arrived earlier. It is easy to see that, for any arrival order, the total surplus generated by L gets exactly allocated to the set of all agents. The Myerson value then allocates the average incremental contribution of a player to the total surplus, taken over all possible orders of arrivals (permutations) of the players, as the player's share of the total surplus. Thus, agent i's Myerson value is[21]

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

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

To start with, consider a procedure proposed in 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

---

[20] See, for example, Arrow (1965).

[21] Our assumption that there is perfect risk sharing among path-connected agents ensures that a coalition of path connected agents generates the same surplus regardless of the exact network structure connecting them. This means that we are in the communication game world originally envisaged by Myerson. We do not require the generalization of the Myerson value to network games proposed in Jackson and Wolinsky (1996), which somewhat confusingly is also commonly referred to as the Myerson value. Jackson (2005) critiques the generalization of the Myerson value on the grounds that the allocation is insensitive to heterogeneities in people's surplus-generating capabilities that are captured by alternative, unformed networks. These heterogeneities are not possible in communication games.

paper, such a procedure leads to the splitting of the total endowment on 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 robust to renegotiation.[22]

Formally, for network L, let $u_i(L)$ be i's expected payoff ex ante (before incomes are realized). We assume that for every $l_{ij} \in L$, agents i and j meet to negotiate a transfer before the endowments are realized. We let each agent have the option of holding up the other by suspending the link. We assume that when negotiating about such suspended links, the agents "split the difference" and benefit equally from the link. Robustness to these renegotiations then requires that, at the margin, each formed link benefits both agents equally. Of course, in order to calculate what agents i and j would receive in the network without the link l_{ij}, we have to consider what would happen if links were renegotiated in the network without l_{ij} and so on. The result is a recursive system of equations. The value of the link is directly pinned down only 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 $\Re^{|\mathbf{N}|}$, representing payoff vectors to agents for different network realizations. We refer to elements of $U(L)$ as contingent payoff schemes given L. For a contingent payoff scheme u, let $u_i(L')$ denote the payoff of agent i given $L' \subseteq L$. Lastly, for ease of exposition, we use $L - l_{ij}$ instead of $L \backslash \{l_{ij}\}$ for the network obtained from L by deleting link l_{ij}.

Definition 2. *For any network L, the payoff vector $(u_1, ..., u_{|\mathbf{N}|})$ is* robust to renegotiation *if there is a contingent payoff scheme u given L such that $u_i = u_i(L)$ for every $i \in \mathbf{N}$ and the following conditions hold:*

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

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

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

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

[22] Such an approach is also reminiscent of Bennett (1988). However, she works in a different environment (the marriage market) and does not select the Myerson value.

Proof. We will use the axiomatization of the Myerson value established in Myerson (1980).[23] This axiomatization states that there exists a unique payoff division rule satisfying the conditions that (i) any link benefits the two connecting agents equally, and (ii) the outcome is efficient at the component level. Formally, these requirements are equivalent to properties (i) and (ii) in Definition 2, so a payoff vector is *robust to renegotiation* if and only if it corresponds to the Myerson value. Since the Myerson value is unique, there is a also a unique payoff vector that is *robust to renegotiation.* □

Proposition 3 is a direct implication of Myerson's axiomatization of the value. A special case of Proposition 3 is Theorem 1 of Stole and Zwiebel (1996), which restricts attention to a star network. Our contribution is to point out that the connection between robustness to renegotiation and the Myerson value applies across all networks.

3. Investing in Social Relationships

Having defined how formed networks map into risk-sharing arrangements, we can now consider agents' incentives to make social investments. We begin by providing the overall framework for the analysis. Then we look at a special case of our model, in which there is a single group. Building on these results we then consider the multiple group case.

In this section we formalize a game of network formation in which establishing links is costly, define efficient networks and identify different types of investment inefficiency.

We consider a two-period model in which in period 1 all agents simultaneously choose which other agents they would like to form links with, and in period 2 agents agree upon the ex-ante Pareto-efficient risk-sharing agreement specified in the previous section (i.e., the total surplus from risk sharing is distributed according to the Myerson value), for the network formed in the first period.[24]

Formally, in period 1 we consider a network formation game along the lines of Myerson (1991): all agents simultaneously choose a subset of the other agents, indicating who they would like to form links (relationships) with. A link is formed between two agents if and only if they both want to form it (i.e., if both agents select each other). When agent i forms a link, he pays a cost $\kappa_w > 0$ if the link is with someone in the same group and $\kappa_a > \kappa_w$ if the link is with someone from a different group. This specification assumes that two agents forming a link have to pay the same cost for establishing the link. However, all of our results below would remain valid if we allowed the agents to share the total costs of establishing a link arbitrarily (namely, if we allowed the agents in the first period not only to indicate who they would like to establish links with, but to also propose a division of the costs of

[23] An assumption made by Myerson when defining the Myerson value is that the surplus generated by a connected component is independent of the network structure within the component. Networks are viewed as communication structures.

[24] For a complementary treatment of network formation when surplus is split according to the Myerson value, see Pin (2011).

establishing each link; a link would then form only 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 and 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 autarky 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 $\{u_i(L)\}_{i\in\mathbf{N}}$ if and only if for all $i,j \in \mathbf{N}$, (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 network 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 4 (Jackson, 2003)**.** *There exists an equilibrium in the game of network formation.*

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

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

where, for $L' \subset L$, $\Delta\,\mathrm{Var}(L,L')$ 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.[25]

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

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

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

4. Local Network Formation

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

The social value of a non-essential link is 0. We begin the analysis by characterizing the social value of an essential link. As shown in the previous subsection, the social value of a link is proportional to the reduction it implies, through a Pareto-efficient risk-sharing agreement, in the sum of the consumption variances. Let $L(\mathbf{S}_1)$ and $L(\mathbf{S}_2)$ be the network components of agent i and agent j on network $L \setminus \{l_{ij}\}$, 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 $\frac{s_2+s_2(s_2-1)\rho_w}{s_2}\sigma^2$, respectively. 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 $\Delta\,\mathrm{Var}(L\cup\{l_{ij}\}, L) = (1-\rho_w)\sigma^2$. This means that, in the case of homogeneous agents, the variance reduction—and therefore the surplus created by an essential link—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 in $\mathbf{S}_2$ when they get 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 in $\mathbf{S}_1$ 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 homogeneous agents these two effects perfectly cancel each other out.

[25]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 the net total surplus is possible by the simultaneous addition and removal of edges).

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

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

Since the surplus created by any essential link is $V \equiv \frac{\lambda}{2}(1-\rho_w)\sigma^2$, the 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 arrival orders of the players. When agent i arrives, the number of component will be reduced by at most 1 for every link formed by i: The link l_{ij} will reduce the number of components in the graph by one when i arrives, if and only if j has already arrived and there is no other path between i and j. For a given arrival order, let $\mathbf{S}$ be the set of agents that arrive before i. 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 arrives 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 and only if there is no other path from i to j on $L(\mathbf{S} \cup i)$. That is, the link l_{ij} is valuable to i when i arrives if and only if it is essential on the graph $L(\mathbf{S} \cup i)$.

Remark 6. *Let $L^e(\mathbf{S}\cup i) \subseteq L(\mathbf{S}\cup i)$ be the set of essential links on $L(\mathbf{S}\cup i)$ and let $L^e_i(\mathbf{S}\cup i) := \{l_{ij} \in L^e(\mathbf{S}\cup i)\}$ be the subset of such links i has formed. The incremental surplus generated when i arrives is proportional to $|L^e_i(\mathbf{S}\cup i)|$. More precisely,*

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

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 of that path. If there are K minimal paths 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. For every $k \in \{1, \ldots, K\}$, let $|P_k(i,j,L)|$ be the cardinality of the set of agents on the minimal path $P_k(i,j,L)$.[26] We can now use these definitions to define a quantity that captures how far away two agents are on a network in terms of the probability that for a random arrival order they will be connected without a direct link when the second of the two agents arrives. We will refer to this distance as the agents' Myerson distance:

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

[26]For example, for a path $P_k(i,j,L) = \{i, i', i'', j\}$, $|P_k(i,j,L)| = 4$ and for a path $P_{k'}(i,j,L) = \{i, i', i''', i'''', j\}$, $|P_{k'}(i,j,L)| = 5$. Finally, we we will let $|P_k(i,j,L) \cup P_{k'}(i,j,L)| = 5$ denote number of different agents on path $P_k(i,j,L)$ or path $P_{k'}(i,j,L)$.

This expression calculates the probability that for a random arrival order the link l_{ij} will be essential immediately after i arrives, using the classic *inclusion–exclusion principle* from combinatorics. This probability is important because it affects i's incentives to link to j.

As an illustration, suppose that there is a unique indirect path $P_1(i,j,L)$ between i and j that contains K agents, including i and j. We then have $md(i,j,L) = 1/2 - 1/K$. To see where this expression comes from, note that there are two reasons why l_{ij} might not be essential when i arrives. First, j might not yet have arrived. This occurs with probability 1/2. Second, all other agents on the path $P_1(i,j,L)$, including j, might have already arrived. This occurs with probability $1/K$. 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 arrives, is $1 - 1/2 - 1/K = md(i,j,L)$.

Suppose now that there are two (minimal) 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 $P_2(i,j,L) = \{i, i', i''', i'''', j\}$. We need to find the probability that all the agents on at least one of these paths are present when i arrives. To avoid double counting, we need to add the probability that all the agents on $P_1(i,j,L)$ are present (1/4) to the probability all the agents on $P_2(i,j,L)$ are present (1/5) and then subtract the probability that all the agents on both paths are present (1/6).[27] So $md(i,j,L) = 1 - 1/2 - 1/4 - 1/5 + 1/6$. The Myerson distance calculation provides the general way of accounting for the probability that at all the agents on at least one of the possible paths are present.

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

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

Proof. See Appendix A. □

The first step in the proof of Proposition 7 establishes that the value of a link l_{ij} to i (and j) is V if the link is essential when i arrives and 0 otherwise. Suppose a link l_{ij} is essential on L. Then for any arrival order, there will always be a component reduction of 1. Therefore, $md(i,j,L) = 1/2$, and l_{ij} will be formed as long as $V > 2\kappa_w$. As V is the social value of forming the link and $2\kappa_w$ is the total cost of forming it, with homogeneous agents there is never underinvestment in equilibrium. This argument is formalized in Proposition 8.

Proposition 8. *If all agents are homogeneous then there is never underinvestment in equilibrium. Furthermore, there is never overinvestment in an essential link.*

[27]Note that all the agents on both $P_1(i,j,L)$ and $P_2(i,j,L)$ will be present if and only if i', i'', i''', i'''' and j are present before i arrives.

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 greater than the overall cost of forming the link, so that $TS(L \cup l_{ij}) - TS(L) > 2\kappa_w$.

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

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

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

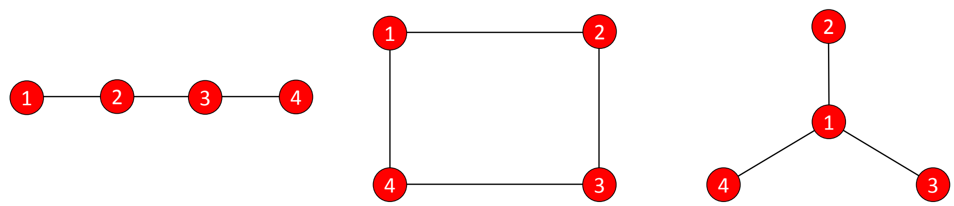
Corollary 10. *For homogeneous agents, in any efficient equilibrium $u_i(L) = |\mathbf{N}(i; L)|(V/2 - \kappa_w)$ and agents' payoffs are proportional to their degree centralities.*

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

Although, as shown above, with homogeneous 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 values of the participants, through increasing their incremental contributions for some arrival orders (for example when i and j are the first two agents to arrive).

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 payoffs implied by L.

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



(A) Efficient, unstable

(B) Stable, inefficient

(C) Efficient, stable

FIGURE 1. Stable and efficient networks for $2\kappa_w \in (\frac{2}{3}V, V)$, where V is the social value of an essential link

Let the line network be the unique (tree) network, up to a relabeling of agents, in which there is a path from one (end) agent to the other (end) agent that passes through all other agents exactly once (see Figure 1a). Let the star network be the unique tree network, up to a relabeling of agents, in which one (center) agent is connected to all other agents (see Figure 1c). On all tree networks connecting at least three agents, there are some agents who have degree 1 (leaf nodes) and some agents who 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 depends only on the length of the path connecting them, in a strictly increasing manner. Suppose d is the number of agents on the unique path connecting i and j. The probability that this path exists when agent i arrives is $1/d$. In addition, if agent j has not yet arrived, which occurs with probability $1/2$, i would not benefit from the link l_{ij}, so i's expected payoff from forming a superfluous link to j is $(1 - 1/2 - 1/d)V$. As d gets large, this converges to $V/2$ which is the value i receives from forming an essential link. These claims are formalized in the next proposition.

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

Proposition 11. *For homogeneous agents, if L is efficient then the following hold:*

(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 $\overline{d}(\kappa_w, V)$, where $\overline{d}(\kappa_w, V)$ is increasing in κ_w, decreasing in V and integer valued.*

Proof. Recall that for a tree network L with diameter $d(L)$, there exist 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 Proposition 7, i and j will want to form the link if and only if

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

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

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

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

Setting $\overline{d}(2\kappa_w) = \lceil 2V/(V - 2\kappa_w)\rceil$ completes the proof. □

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

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

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

5. Connections Across Groups

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

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

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

Recalling that

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

some algebra yields[28]

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

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

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

Motivated by Proposition 13, within a group we will continue to focus on the problem of overinvestment rather than underinvestment. However, in contrast to Proposition 13, there can be underinvestment across groups. The key insight is that, as opposed to the case of homogeneous agents, where the value of an essential link does not depend on the sizes of the components it connects, the value of an essential link connecting two different groups of agents increases in the sizes of the components. To demonstrate this formally, consider an

---

[28]One of the key steps to simplifying the expression is noting that $2\sum\limits_{i=0}^{k-1}(s_i \sum\limits_{j=i+1}^{k} s_j) = \left(\sum\limits_{i=0}^{k} s_i\right)^2 - \sum\limits_{i=0}^{k} s_i^2$.

isolated group that has no across-group connections and consider the incentives for a first such connection to be formed. Let the first component consist of agents from a single group, say group 0, and let the second component consist of agents from one or more of the other groups (1 to k). The variance reduction obtained by connecting these two components then simplifies to

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

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

The inequality follows since $\rho_w > \rho_a$.

An immediate implication is that agents i and j who connect two otherwise unconnected groups receive a strictly smaller combined private benefit than the social value of the link. To see why, suppose that on the network L the link l_{ij} is essential, and without l_{ij} there would be two components, the first connecting agents from group $G(i)$ and the second connecting agents from group $G(j) \neq G(i)$. Consider the Myerson value calculation. For arrival orders in which i or j is last to arrive, the value of the additional variance reduction due to l_{ij} obtained upon the arrival of the later of i or j, is the same as its marginal social value, i.e., the value of variance reduction obtained by l_{ij} on L. For any other arrival order the value of variance reduction due to l_{ij} when the later of i or j arrives is strictly less. Averaging over these arrival orders, the link l_{ij} contributes less to i and j's combined Myerson values than its social value, leading to the possibility of underinvestment.

We formalize the resulting possibility of underinvestment in Proposition 14.

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

Proposition 14. *If $|\mathbf{M}| \geq 2$, then underinvestment is possible in equilibrium.*

Proof. We will show that if there are $|\mathbf{M}| \geq 2$ 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 formation of an extra link connecting any two groups having a strictly positive social value. By assumption, the within-group correlation coefficient $\rho_w < 1$ and the coefficient of absolute risk aversion $\lambda > 0$. Together, these parameter restrictions imply that the certainty-equivalent value of a variance reduction from connecting one agent to any group of other agents is strictly positive. Thus, for κ_w sufficiently close to 0, in all equilibria any two agents from the same group must be path connected.

Assume now that all groups form separate network components, and consider a potential extra link l_{ij} connecting groups g and g'. As shown above, the change in total variance, and so the surplus, achieved by connecting agents in $\mathbf{S}_g$ to agents in $\mathbf{S}_{g'}$ is increasing in the sizes of both groups, s_g and $s_{g'}$ respectively. This means that the marginal contribution of the link l_{ij} to total surplus is greater than its contribution to total surplus when the later of i and j arrive, unless i or j arrives 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 $j \in \mathbf{S}_{g'}$. There thus exists 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 agents are completely connected within groups, but there is no link across groups, despite the formation of such a link being socially desirable. □

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

We begin by considering *local overinvestment* within groups, which corresponds to the formation of superfluous within-group links. We found in the previous section that for homogeneous agents, the star was the efficient network that minimized the incentives for overinvestment. However, once we include links to other groups, the analysis is more complicated. The variance reduction a within-group link generates is still 0 if the link is superfluous, but when the link is essential it 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 numbers of people in those groups.[29] This makes Myerson value calculation substantially more complicated. With homogeneous 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 that are being connected. Nevertheless, our earlier result generalizes to this setting, although the argument establishing the result is more subtle.

[29]In the case of an essential across-group link that bridges two otherwise disconnected groups, the comparative statics are unambiguous. In this case, the variance reduction is increasing in the sizes of the groups connected, as shown by inequality (4).

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

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

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

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

We now consider the local network structures that maximize the incentives for an across-group link to be formed. We have already established that the marginal contribution of a first bridging link to the total surplus is increasing in the sizes of the groups it connects. By the Myerson calculation, the agents with the strongest incentives to form such links are then those who will be linked to the greatest number of other agents within their group when they arrive. The result below formalizes this intuition.

Let $\mathcal{A}(\mathbf{S}_k)$ be the set of possible arrival orders for the agents in $\mathbf{S}_k$. For any arrival order $A \in \mathcal{A}(\mathbf{S})$, let $\mathbf{T}_i(A)$ be the set of agents to whom i is path connected on $L(\mathbf{S}')$, where $\mathbf{S}'$ is the set of agents (including i) that arrive weakly before i. Let $T_i^{(m)}$ be a random variable, taking values equal to the cardinality of $\mathbf{T}_i(A)$, where A is selected uniformly at random from those arrival orders in which i is the m-th agent to arrive.

We will say that agent $i \in \mathbf{S}_k$ is more central within his group than agent $j \in \mathbf{S}_k$ 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 arrival orders in which i is the m-th agent to arrive, and all the arrival orders in which j is the m-th agent to arrive, 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. [30] This measure of centrality provides a partial ordering of agents.

Proposition 16. *Suppose agents in $\mathbf{S}_0$ form a network component, and all other agents in $\mathbf{N}$ form another 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 l_{ij} than i' receives from forming $l_{i'j}$:*

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

Proof. See Appendix A. □

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

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

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

Proof. Within a local network, an agent connected to all other agents is weakly more central than any other agent. For any arrival order, such an agent is always connected to all other agents that arrive before him. As no agent can be connected, at the time of his arrival, to an agent who arrives later, an agent connected to everyone else is more central than all other agents—and so, by Proposition 16, has stronger incentives to form an across-group link.

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

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

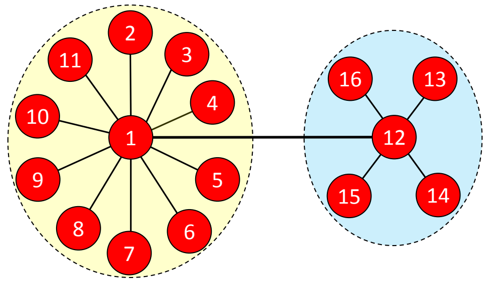


FIGURE 2. Center-connected local stars, in a context with two groups.

The above establishes that for two groups, the efficient networks that minimize within-group overinvestment and across-group underinvestment 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 within-group overinvestment, it also minimizes the incentives for across-group underinvestment. If an agent i provides an across-group link, then of all the possible local network structures, the star with i at the center maximizes i's incremental payoff from establishing the link.[31]

6. Empirical Analysis

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

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

6.1. Setting and Data.

The data we use were collected in 2011 by Banerjee, Chandrasekhar, Duflo and Jackson (2013a, 2014a), by conducting surveys in 75 villages in Karnataka, India. The villages span 5 districts and range from 2- to 3-hour drive from Bangalore. They are far

[31]Nevertheless, for some (but not all) parameter values, a local star will be more equitable when the central agent forms an across-group link than without it, because the across-group link generates positive spillovers to the whole group.

enough apart to be treated as independent systems (the median distance between them is 46 km, and a district has 1000–3000 villages). The survey included a village questionnaire, a census of all households, demographic covariates (including caste and occupation), as well as data on a number of amenities (e.g., roofing, latrine, or electricity access quality). A detailed individual-level survey was administered to most adults in every village. The survey included a networks module with twelve dimensions of relationships, including financial relationships, social relationships, and advice relationships.[32]

Our analysis focuses on two types of networks: the *financial graph*, L^F, and the *social graph*, L^S. The financial graphs represent risk-sharing connections, and the social graph represents friendships/links used to socialize. We build "AND" networks, which say a link exists if it exists on every dimension being considered (various types of financial connections on the financial network, and various types of social connections on the social network).[33] The advantage of doing this is that it generates a network structure that is more robust to independent measurement error.[34] The event that a frivolous link is coded decreases exponentially if we require that it exist across multiple dimensions, which a priori helps in detecting effects when we look differentially across network-type. In some of our empirical analysis, we will explicitly consider how our predictions differentially play out in L^F relative to L^S, as our theory speaks to the former.

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

Table 1 provides some summary statistics for our data. The average number of households per village is 209 with a standard deviation of 80. The average degree of the financial graph is 3.1, and of the social graph is 1.7. It is interesting to note that the financial graph is more dense. Further, we see that the clustering in the financial graph is 0.18 (0.07), whereas for the social graph it is 0.05 (0.03). Though the density is only 1.8 times as high for the financial graph than for the social graph, the fact that the clustering is over 3 times as high for the financial graph is of note. This is consistent with the results of, for instance, Jackson et al. (2012) that financial links may need to be supported/embedded in cliques to sustain cooperation. Both the financial and social networks exhibit relatively few cross-caste links. This is seen by looking at the ratio of the probability of having a cross-caste link relative to the probability of having a within-caste link. A within-caste link is five times as likely in the

[32]See Banerjee et al. (2014a) for more details. In total we have network data from 89.14% of the 16,476 households based on interviews with 65% of all individuals of age 18 to 55.

[33]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.

[34]We say $l_{ij} \in L^S$ if i goes to j's house to socialize or vice versa.

financial graph and three times as likely in the social graph. Finally, 67% of households are high caste (GM or OBC). We describe our main outcome variables later.

6.2. **Empirical Framework.**

6.2.1. *Predictions.* The broad predictions of our model are that (1) there is endogenous centrality, (2) there is no underinvestment within groups, (3) agents cannot be too far away from each other (in terms of the Myerson distance), (4) agents that have across-caste links should be more Myerson central. However, these predictions do not provide a clean test of our theory. Predictions (2) and (3) depend on an unobservable linking cost parameter and there are many alternative stories, including ones not directly connected to risk sharing, that generate similar predictions. For example, if individuals have heterogenous time budgets and made random links within and across groups, predictions (1) and (4) are mechanically generated. We therefore turn to more subtle predictions that rely on the comparative statics under our model as we change parameters of the economic environment. We study four demanding predictions from the theory with richer empirical content.

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

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

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

$$md(i,j,L) \leq \frac{2\kappa_w}{(1-\rho_w)\lambda\sigma^2}. \tag{5}$$

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

Consider the variance reduction obtained by combining any two components, with agents in groups $\mathbf{S}_0, \ldots, \mathbf{S}_k$ and $\widehat{\mathbf{S}}_0, \ldots, \widehat{\mathbf{S}}_k$ respectively. From equation 2, for some function f, the certainty-equivalent value of this variance reduction is

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

This is hard to compute in general. We approximate it by assuming that within each component there are the same number of people from each group: $s_i = \alpha$ and $\hat{s}_i = \beta$ for $i = 0, \ldots, k$.

Proposition 18. *The certainty-equivalent value of variance reduction obtained by linking a component with α people from each of groups $0, ..., k$ to a component with β people from each of groups $0, ..., k$ is*

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

Proof. See Appendix A. □

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

The above considerations lead to the following predictions:

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

P2. In villages with lower ρ_w, the average Myerson distance should be smaller.

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

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

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:

P3. In villages with higher σ^2, the association between within-caste centrality and the formation of across-caste links is lower.

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

6.2.2. *Empirical strategy.* Our predictions are about how network structure varies with σ^2, ρ_w, and ρ_a. The analysis is observational (not causal) and simply looks at the cross-sectional variation of network structure with these parameters through ordinary least squares (OLS).

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

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

At the same time, we note that our empirical approach is more conservative than similar studies in the literature (e.g., Karlan et al. (2009), Ambrus et al. (2014), Kinnan and Townsend (2012)) in terms of statistical inference. First, these studies typically have very few networks (2, 1, and 16, respectively), and therefore consider node or link-level regressions with standard errors generated at that level. This effectively treats nodes or dyads as independent or loosely correlated, making inference preclude, essentially, village-level shocks. Correlation in any factors outside the model (e.g., incentives to form links for other reasons) as well as equilibrium selection at the village level are all precluded from econometric analyses that don't study the theory where the entire graph is the unit of observation. We make no such assumptions on the independence of nodes or dyads for valid statistical inference, and instead exploit the fact that we have 75 independent villages (recall that the median pairwise distance between them is over 46 km). By focusing on village-level variation, we are allowing for arbitrary correlation within graphs. In fact, in our most conservative specifications, we allow correlation at the subdistrict level.

Second, analyses in the previous literature usually do not have access to different types of edges—the multigraph—and therefore cannot employ our difference-in-differences approach.

What we do, relative to this, is extremely conservative. By differencing across network types, we are asking whether the patterns in the graph which match our theory are differentially at play for the financial network relative to the social networks.

Third, another reason our approach is conservative is that typical models of multigraph link formation have a fixed-cost component. Thus, incentives driving the formation of risk-sharing links are likely to influence the structure of social or information links through this channel. Observe that by looking at the financial graph relative to the social graph, the variation coming from the fixed-cost component – which is consistent with our theory – is not even being used in our analysis in support of our theory.

6.3. **Variable Construction.**

6.3.1. *Approximating the Myerson distance and centrality.* Our next task is to compute the Myerson distance of every pair in every village and the Myerson centrality for all nodes. Unfortunately, this is computationally infeasible for the sample sizes of our data (see Algaba et al. (2007)), presenting a new challenge. Thus, we develop an approximation, described below.

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 his arrival agent i will not be connected to agent j. It is difficult to directly characterize $\mathbf{md}(L)$ (or equivalently, $\mathbf{q}(L)$) as each village typically consists of around 230 households and the number of candidate paths between each i and j is exponential in the size of the network. Correctly accounting for paths that share nodes is computationally very intensive (see Proposition 7), and it has to be done for all pairs of agents without a direct connection.[35] Instead, we develop a computationally feasible approximation of $\mathbf{md}(L)$, which is exact for trees.

To approximate $\mathbf{q}$, we use the following idea. The algorithm works by starting with a node, moving to its neighbors, then move to its neighbors' neighbors, and so on, never returning to a previously used node along a given path. This helps us to avoid counting non-minimal paths. All the while, we keep track of how many ways we have moved from the original node to any given node. We denote our approximation of $\mathbf{q}$ by $\widehat{\mathbf{q}}$.

The inclusion–exclusion principle weights paths that are longer less and a path that shares many nodes with another less. With this in mind, we make the following two approximations. Let the shortest path between two nodes be of length l. We first count the paths of length l and length $l+1$. We then count paths of length $l+2$.[36] If there are fewer than k such paths, we use them all. Otherwise, we consider only the k shortest and in practice we set $k=4$.[37] Discarding longer paths in this way biases downwards our approximation of $\mathbf{q}$. As

[35]Further, due to presumed measurement error (see Banerjee et al. (2013)), there are likely to be missing paths. In fact, the data have occasional disconnected components, and so measures that are precisely based on exact paths or even maximal path lengths are likely to be problematic (Chandrasekhar and Lewis (2014)).

[36]Counting more paths greatly (exponentially) increases the running time of our algorithm.

[37]We need a fixed (small) truncation. Otherwise both the memory requirements and the run-time of the algorithm grow exponentially. Results are not sensitive to the truncation point.

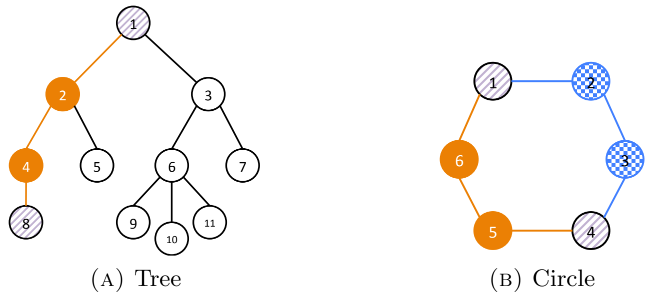


FIGURE 3. The nodes i,j for which we are computing $md(i, j, L)$ have purple stripes. The tree contains a single minimal path (solid orange nodes), whereas the circle contains two paths (solid orange nodes and chequered blue nodes).

we cannot keep track of exactly which nodes feature in each path, we also have to make an assumption about the overlap of nodes in order to apply the inclusion–exclusion principle to these paths. Each path must share the same first and last node. We perform the inclusion–exclusion principle assuming that only these nodes are shared (see Section 4). Assuming no other nodes are shared introduces a second bias, but this time upwards in our approximation.

To explain these concepts, we provide some illustrations. Figure 3 presents two examples: a tree and a circle. The tree has a single minimal path between nodes 1 and 8, whereas the circle has two minimal paths between nodes 1 and 4. Figure 4 shows how links are removed for the case of a tree. Once a node has been reached, links back into that node are deleted before the nodes neighbors are "infected." This ensures only minimal paths are included in the calculation.

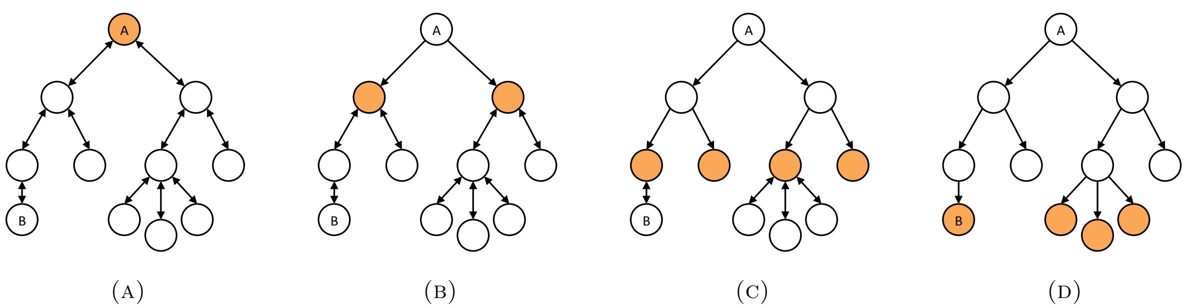


FIGURE 4. As the algorithm progresses, directed links into nodes that are reached are deleted. This ensures that only minimal paths are included. In this case, as in all tree networks, there is a unique minimal path from A to B.

In the case of the circle shown in Figure 3, our algorithm is also exact for paths between 1 and 4. There are two minimal paths (which in this case are both shortest paths too), and we find both in the initial run of our algorithm. Following the inclusion–exclusion principle, we add 1/4 to 1/4 and subtract 1/6. In this case our assumption that the two paths share only two nodes is accurate. We are also exact for paths between 1 and 3, but in this case there is a path of length $l+2$. To find this path, we look for paths of length l from 1 to nodes other than 3. In this case there is one such path to node 5. We then look for paths from 5 to 3 that pass through one other node. There is one such path and so the calculation we perform is: $1/3+1/5-1/6$. While we are accurate for all pairs of nodes in the circle shown, in larger circles we will miss the longer paths.

The following algorithm finds the length of the shortest path between two nodes, how many paths of that length there are and how many paths there are that are one longer. From this information, we also find paths of length $l+2$.[38]

Algorithm 19 (Incoming Link Deletion)**.** *Let e^i be the ith basis vector. This will represent the root (starting) node. Initialize $\widehat{\mathbf{q}} = zeros(n,n)$, a matrix of zeros. Initialize $z^{t,i} = zeros(n,1)$ and $x^{t,i} = zeros(n,1)$ to be n-vectors of zeros, indexed by $i=1,...,n$ and $t=1,...,T$. Repeat steps 1–4 for each of $(e^1,...,e^n)$.*

(1) *Period 1: There is no identification or updating steps.*
  (a) *Percolation: $x^{1,i} = \mathbf{A}e^i$.*
  (Identifies who is connected to the root node)
(2) *Period 2, given $(x^{1,i}, \mathbf{A})$:*
  (a) *Identification: $z^{2,i} = e^i$.*
  (b) *Update graph:*[39] $\mathbf{A}_2 = zeros(n,n),\ \mathbf{A}_2(\neg z^{2,i}, :) = \mathbf{A}(\neg z^{2,i}, :)$.
  (Deletes links into the root node)
  (c) *Percolation: $x^{2,i} = \mathbf{A}_2 x^{1,i}$.*
  (Records number of paths from root node to other nodes passing through one other)
(3) *Period t, given $(x^{t-1,i}, \mathbf{A}_{t-1})$:*
  (a) *Identification: $z^{t,i} = \mathbf{1}\left\{\sum_{s=3}^{t} x^{s-2,i} > 0\right\}$.*
  (Identifies nodes already visited)
  (b) *Update graph: $\mathbf{A}_t = zeros(n,n),\ \mathbf{A}_t(\neg z^{t,i}, :) = \mathbf{A}_{t-1}(\neg z^{t,i}, :)$.*
  (Deletes links into all nodes that have already been visited)
  (c) *Percolation: $x^{t,i} = \mathbf{A}_t x^{t-1,i}$.*

By construction $x_j^{t,i}$, the jth entry of $x^{t,i}$, records paths from i to j that pass through t nodes. If t' is the lowest t with a positive entry in this matrix, then the shortest path from i to j passes through t' nodes. In this case, $x_j^{t',i}$ tells us how many such paths there are and $x_j^{t'+1,i}$ tells us how many paths there are that pass through one more node. However, by construction $x_j^{t'+k,i} = 0$ for all $k > 1$ and longer paths are not recorded. This is because the incoming links to node j will have been deleted by this step of the algorithm. Deletion of

[38]For paths from i to j, this is done by looking at paths of length l to agents other than j, and then looking at paths from these agents to j.

[39]Let $\mathbf{A}(:,v)$ denote $(A(1,j),...,A(n,j))$.

incoming links helps prevents non minimal paths from being recorded. Using this information for all seed nodes, the number of paths of length $t' + 2$ between i and j are also found as described above. The inclusion-exclusion principle is then applied to this combined set of paths, assuming each path shares only the first and last nodes, to calculate $\widehat{\mathbf{q}}(L)$.

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

Proof. See the Appendix A. □

To operationalize $\widehat{\mathbf{q}}(L)$ in our regression analysis, we need a village-level measure of Myerson distances. We use $\widetilde{q}(L) := \sum_{i<j} \widehat{q}_{ij} / \binom{n}{2}$ which measures an appropriately weighted density of the network. Finally, to approximate Myerson centrality we use $\sum_j \widehat{q}_{ij}$, as people are central when they are likely to be connected to others. Thus, their q_i terms are high (equivalently, they have a low Myerson distance to others).

A limitation of the Incoming Link Deletion algorithm is that longer paths are excluded. To address this, we construct an alternative algorithm. This Outgoing Link Deletion algorithm is identical to the one described, except that it deletes outgoing links instead of incoming links. The Outgoing Link Deletion algorithm finds longer paths, and does an especially good job of picking up longer paths that share few nodes with other paths. However, it also includes additional short non-minimal paths and is not exact for tree networks. As longer paths are found, we directly use the output of the algorithm without constructing any additional longer paths. Nevertheless, for the set of paths we find, it is computationally infeasible to compute the Myerson distances using the inclusion-exclusion principle. Censoring these paths would defeat the point of the Outgoing Link Deletion algorithm. Instead, we use an approximation of the inclusion–exclusion principle which makes the computation much simpler. This approximation treats every path as completely independent, assuming that no nodes are shared (even though we know at two must be). For example, if we find 3 paths from i to j that pass through l nodes, l' nodes and l'' nodes respectively, our approximation of q_{ij} will be $1/l + 1/l' + 1/l''$.

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

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

Measuring income correlation is difficult. Ideally, we would have time-series data on incomes of all households, as well as plausible instruments allowing us to calculate the exogenous variation in income correlation both within and across castes for each village. While we do not have access to income data, we do have detailed data on occupation.[40] Thus, we make use of the relative within-caste to across-caste occupation correlation. The main idea is that shocks to individuals will be more highly correlated when they have the same occupation. We take two approaches to computing caste–income correlation.

One approach is to look at the correlation of being in the high-caste group with holding a given occupation, for all occupations in our survey. We then take the weighted average of these correlations, where the weight is the share of agents in the occupation. Thus, for a given village we consider

$$\widehat{\rho_w - \rho_a} := \sum_{k=1}^{K} \operatorname{corr}\left(Caste, Occupation_k\right) \Pr\left(Occupation_k\right),$$

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

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

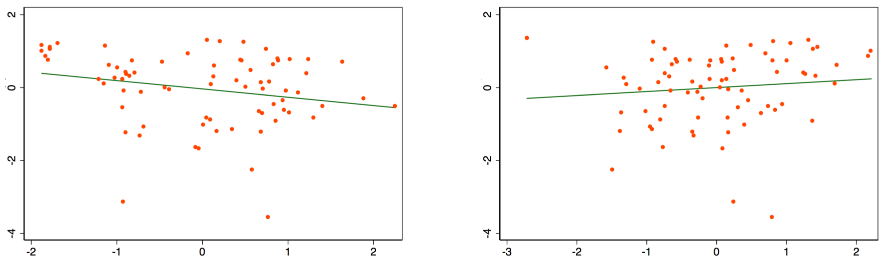
$$\widehat{\rho_w - \rho_a} := \sum_{g} \phi_g \Pr(o_{i,g} = o_{j,g}) - \Pr(o_{i,A} = o_{j,B}),$$

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

(1) $\widehat{\rho_w - \rho_a}^{I} = \sum_g \phi_g \Pr(o_{i,g} = o_{j,g}) - \Pr(o_{i,A} = o_{j,B})$.
(2) $\widehat{\rho_w - \rho_a}^{II} = \sum_{k=1}^{K} \operatorname{corr}\left(Caste, Occupation_k\right) \Pr\left(Occupation_k\right)$.

6.3.3. *Myerson distance as a function of income variability and correlation.* We begin with P1 and P2. Figure 5 presents results in the raw data. As predicted, villages variable individual incomes are associated with lower $\widetilde{md}(L^F)$, and villages with more within-caste income correlation are associated with higher $\widetilde{md}(L^F)$.

[40]We recognize that occupations can have a choice component. Nevertheless, we proceed with these measures for three reasons. First, in rural villages the primary household occupation (agriculture or sericulture) is often passed on through generations. Second, it is possible to show that under a natural model with endogenous selection into occupation, the within- versus across-group income correlations are captured by our occupational choice measures, whereas the choice of occupation does not generate spurious correlations with the network in a manner consistent with P1–P4. Finally, this is the best possible approximation, given the severe income data limitations, as a necessary component of our analysis is the network data.



(A) Measure: $\widetilde{md}(L^F)$ vs σ^2 (B) Measure: $\widetilde{md}(L^F)$ vs ρ_w

FIGURE 5. Myerson distance in standardized units against the variability of rainfall at the village level and the within-caste income correlation metric, both in standardized units.

Table 2 demonstrates the robustness of this graphical evidence in regression of $\widetilde{md}(L^F)$ on σ, ρ_w, and with district or subdistrict fixed effects as well as controls for caste composition. In all specifications here and throughout the paper, unless otherwise noted we use a Wild clustered bootstrap to account for subdistrict-level clustering in our inference (Cameron et al., 2008).[41]

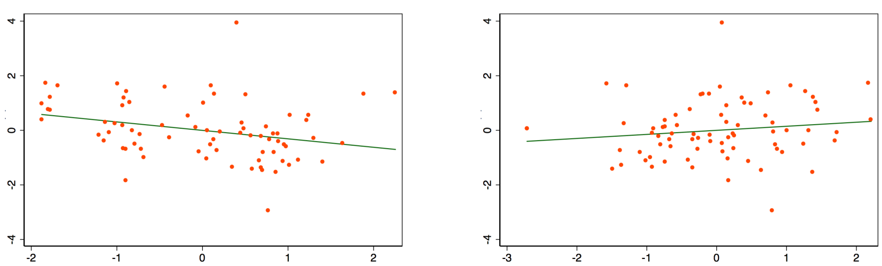
Columns 1-3 present regressions using $\widehat{\rho_w}^I$, whereas columns 4–6 present regressions using $\widehat{\rho_w - \rho_a}^{II}$. Panel A presents results using the Incoming Link Deletion algorithm whereas Panel B looks at Outgoing Link Deletion. We find a one standard deviation increase in Measure I is associated with a 0.292 or 0.296 standard deviation increase in $\widetilde{md}(L^F)$ (column 1, Panels A and B). However, there is no significant relationship between income variability and $\widetilde{md}(L^F)$ in these specifications. Column 4 presents the analogous results using $\widehat{\rho_w - \rho_a}^{II}$. A one-standard-deviation in Measure II is associated with a 0.205 or 0.199 standard deviation increase in $\widetilde{md}(L^F)$, but income variability has no detectable correlation with the outcome variable (column 4, Panels A and B). In columns 2 and 5 we add district fixed effects, and the results are mostly reflective of those in columns 1 and 3. A one standard deviation increase in income correlation corresponds to a roughly 0.1 or 0.18 standard deviation increase in Myerson distance, using Measure I (p-values 0.477 and 0.12 in Panels A and B) and Measure II (p-values 0.00 and 0.012 in Panels A and B) respectively. Inclusion of district fixed effects resolves some of the noise around the rainfall variability coefficient as well in the case of Panel A. Finally, when we include subdistrict fixed effects, our estimate under Measure I is too noisy to distinguish from zero, though we find similar evidence under Measure II (a 0.1 standard deviation effect, p-value 0.04 and 0.13 in Panels A and B). Further, when comparing villages within subdistricts we are able to identify an income variability effect. A one standard

[41]While these villages are essentially independent units, the median distance being 46 km apart, we take a conservative approach because geography is a determinant of rainfall variability and occupation. Our villagers are members of 12 subdistricts, and we therefore cluster our standard errors at that level. To deal with the finite sample bias, we use a Wild cluster bootstrap procedure for the t-test statistic, using Rademacher weights, to generate p values for hypothesis testing as well as standard errors.

deviation increase in income-variability is associated with a 0.3 standard deviation decrease in the average Myerson distance (columns 3 and 6, respectively, of both panels).

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

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



(A) $\widetilde{md}(L^F) - \widetilde{md}(L^S)$ vs σ^2 (B) $\widetilde{md}(L^F) - \widetilde{md}(L^S)$ vs ρ_w

FIGURE 6. Myerson distance, differenced across network type, versus income variability and within-caste income correlation

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

$$Y_{v,t} = \alpha + \beta\sigma_v \mathbf{1}_{\{t=F\}} + \gamma\rho_{w,v}\mathbf{1}_{\{t=F\}} + \mu_v + \mu_t + \delta X_v \mathbf{1}_{\{t=F\}} + \epsilon_{v,t},$$

where $Y_{v,t} = \widetilde{md}(L_v^t)$, μ_v is a possibly endogenous village fixed effect, μ_t is a network type fixed effect, and X_v are village-level controls. P1 and P2 correspond to $\beta < 0$ and $\gamma > 0$, since our theory pertains only to risk sharing networks. Thus, our effects should be differentially more predictive for the financial network than for the social network and should remain true when looking at relative effects.

Table 3 presents the results. Panel A presents the Incoming Link Deletion algorithm and Panel B the Outgoing Link Deletion algorithm. We include results for both measures of income correlation. We find that a one standard deviation increase in income variability differentially decreases $\widetilde{md}(L)$ by about 0.24 standard deviations more in financial networks than in social networks (columns 1-4), irrespective of the measure of income correlation used. We are unable to detect a statistically significant association between rainfall variability and the outcomes of interest.

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

6.3.5. *Association between within-caste centrality and cross-caste links.* We now look at how the composition of cross-caste links varies with these parameters. P3 says that in villages with higher variability we should see a greater association between within-caste centrality and having an across-caste link. P4 says that a similar effect occurs when we look at within- versus across-caste income correlation.

We present regressions where we look at the financial network only on σ_v and $\rho_w - \rho_a$ as well as difference-in-difference results. To make this concrete, the latter regression is

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

where $Y_{v,t}$ is the average within-caste centrality of nodes with a cross-caste link in network t. The remaining terms are as before. P3 and P4 imply $\beta < 0$ and $\gamma < 0$.

Table 4 presents the results. Panel A shows regressions using just the financial networks on regressors of interest. Panel B displays results from the difference-in-difference. Columns 1–3 use the Incoming Link Deletion algorithm, while columns 4–6 use the Outgoing Link Deletion algorithm to compute the outcome of interest.

The main result is that we find a one standard deviation increase in the measure of the $\rho_w - \rho_a$, irrespective of which measure is used, corresponds to about a 0.15 standard deviation decline in the average Myerson centrality (within caste) of the individuals with across caste links. Again, however, when we look at rainfall variability, we are unable to confirm or reject our theory, as estimates are very noisy.

Thus, our evidence is partially in support of our theory. We find a negative association between income correlation and the centrality of those with a cross-caste link. However, we are unable to confirm or reject our hypothesis when looking at the variability of income.

6.4. **Discussion.** This section provided suggestive evidence, consistent with our theory, demonstrating the following:

(1) Networks have higher Myerson distances when within-group income correlation is higher.
(2) Mild evidence in favor of, though mostly evidence leaving the reader agnostic as to whether, networks exhibit lower Myerson distance with greater variability in rainfall.
(3) Networks exhibit more negative associations between within-group centrality and across-group linking when within-group income correlation is higher, though we detect no relationship with variability of rainfall.
(4) Results (1)–(3) are robust to a difference-in-differences approach. By differencing out the social graph, we can remove village-fixed endogenous factors and more convincingly argue that the underlying force concerns risk sharing.

Having several distinct predictions is useful, as we are conducting an observational analysis. We find that we cannot reject any of these predictions based on our data, and in fact, given the data limitations we view our results as surprisingly supportive of the theory.

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

Furthermore, our compositional story is new. More generally, P3–P4 are specific to either risk-sharing relationships between castes or differences in income correlations between castes, and so within the literature are unique to our model. These predictions are also more subtle. For instance, higher within-group correlation leading to more sparse risk-sharing networks is not a conclusion of other risk-sharing network formation stories.

We again emphasize that our findings are clearly observational (non-causal) and subject to measurement error. However, given the unique dataset we used—network data for 75 independent villages, including multigraph data that allows 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 types 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 are consistent with our story. Additionally, we are extremely conservative in conducting statistical inference, by relying only on independence across villages (in fact, we are employing an even more conservative approach than that by using a Wild cluster bootstrap at the subdistrict level) and not exploiting any of the within-village observations (which are likely to be correlated).

7. Conclusion

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

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

Using a unique dataset of 75 Indian villages, we find empirical support for our model. This is particularly important because our model abstracts from the more widely studied enforcement role it might play. Although we certainly do not reject the possibility of the network also playing such a role, we see evidence in our data that the network structure affects how the gains from risk sharing are distributed, suggesting that this is an important consideration when studying the network formation problem. In particular, we find that higher income variability is associated with denser social networks (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 could apply in other social contexts too, as long as the economic benefits created by the social network are distributed similarly to the way they are 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 provide a dynamic extension of the analysis that allows for autocorrelation between income realizations.

References

Ahuja, G. (2000): "Collaboration networks, structural holes, and innovation: A longitudinal study," Administrative science quarterly, 45, 425–455.

Algaba, E., JM Bilbao, J. Fernández, N. Jiménez and J. López (2007): "Algorithms for computing the Myerson value by dividends," Discrete Mathematics Research Progress, 1-13.

Ali, N. and D. Miller (2013): "Enforcing cooperation in networked societies," mimeo UC San Diego.

Ali, N. and D. Miller (2013): "Ostracism," mimeo UC San Diego.

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.

Attanasio, O., A. Barr, J. Cardenas, G. Genicot and C. Meghir (2012): “Risk pooling, risk preferences, and social networks,” AEJ - Applied Economics, 4, 134-167.
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 (2013a): “The diffusion of microfinance,” Science, 341, No. 6144 1236498
Banerjee, A., E. Duflo, M. Ghatak, and J. Lafortune (2013b): “Marry for what? Caste and mate selection in modern India.” American Economic Journal: Microeconomics, 5(2): 33-72.
Banerjee, A., A. Chandrasekhar, E. Duflo and M. Jackson (2014a): “Gossip: Identifying central individuals in a social network,” MIT and Stanford working paper.
Banerjee, A., A. Chandrasekhar, E. Duflo and M. Jackson (2014b): “Exposure to formal loans: Changes in social networks,” MIT and Stanford working paper.
Banerjee, A., E. Breza, E. Duflo and C. Kinnan (2014c): “Does microfinance foster business growth? The importance of entrepreneurial heterogeneity,” mimeo MIT.
Bennett, E. (1988): “Consistent Bargaining Conjectures in Marriage and Matching,” Journal of Economic Theory, 45, 392-407.
Bergstrom, T. and H. Varian (1985): “When do markets have transferable utility?,” Journal of Economic Theory, 35, 222-233.
Billand, P., Bravard, C. and S. Sarangi (2012): “Mutual insurance networks and unequal resource sharing in communities,” DIW Berlin Discussion Paper.
Binzel, C., E. Field and R. Pande (2014): “Does the arrival of a formal financial institution alter informal sharing arrangements? Evidence from India,” mimeo Harvard University.
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.
Burt, Y. (1992): “Structural holes,” Academic Press, New York.
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.

Cameron, A., J. Gelbach and D. Miller (2008): "Bootstrap-based improvements for inference with clustered errors, " The Review of Economics and Statistics, 90(3), 414-427.

A. Chandrasekhar and R. Lewis (2014): "Econometrics of sampled networks," mimeo Stanford University.

Cole, H., and N. 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," Journal of Industrial Economics, forthcoming.

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

Goyal, S. and F. Vega-Redondo (2007): "Structural holes in social networks ," Journal of Economic Theory, 137, 460–492.

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. (2005): "Allocation rules for network games," Games and Economic Behavior, 51, 128-154.

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

Jackson, M., T. Rodriguez-Barraquer and X. Tan (2012): "Social capital and social quilts: Network patterns of favor exchange," The American Economic Review 102.5 (2012): 1857-1897.

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

Karlan, D., M. Mobius, T. Rosenblat and A. Szeidl (2009): "Trust and social collateral," Quarterly Journal of Economics, 124, 1307-1361.

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

Kinnan, C. (2011): "Distinguishing barriers to insurance in Thai villages," mimeo Northwestern University.

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.

Matsuura, K. and C.J. Willmott (2012): "Terrestrial precipitation: 1900–2008 gridded monthly time series,"Center for Climatic Research Department of Geography Center for Climatic Research, University of Delaware.

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

Mehra, A., Kilduff, M. and D. Bass (2001): " The social networks of high and low self-monitors: implications for workplace performance,"Admin. Sci. Quart. 46 (2001) 121–146.

M. Morduch (1991): "Consumption smoothing across space: Tests for village-level responses to risk." Harvard University.

M. Morduch (1999): "Between the state and the market: Can informal insurance patch the safety net?" The World Bank Research Observer 14.2, 187–207.

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

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

Myerson, R. (1980): "Conference structures and fair allocation rules," International Journal of Game Theory, 9.3, 169-182.

Myerson, R. (1991): "Game theory: Analysis of conflict, Harvard University Press, Cambridge.

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

Pin, P. (2011): "Eight degrees of separation," Research in Economics, 65, 259-270.

Podolny, J., and J. Baron, (1997): "Resources and relationships: Social networks and mobility in the workplace," American Sociological Review, 673-693.

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

Slikker, M. (2007): "Bidding for surplus in network allocation problems," Journal of Economic Theory, 137, 493-511.

Srinivas, M. (1962): "Caste in modern India and other essays," Bombay, India: Asia Publishing House.

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.

van Lint, J. H., and R. Wilson (2001): "A course in combinatorics," Cambridge university press.

Walker, T. and J. Ryan (1990): "Village and household economics in India's semi-arid tropics," Johns Hopkins University Press, Baltimore.

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

Table 1. Summary Statistics

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

Table 2. Myerson distance vs. income variability and correlation

	Measure I			Measure II		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Incoming Link Deletion Algorithm					
Income Correlation	0.292	0.086	-0.006	0.199	0.180	0.150
	(0.137)	(0.106)	(0.108)	(0.092)	(0.055)	(0.079)
	[0.046]	[0.477]	[0.935]	[0.034]	[0.000]	[0.044]
Income Variability	0.009	-0.217	-0.278	0.015	-0.278	-0.315
	(0.126)	(0.131)	(0.097)	(0.158)	(0.131)	(0.094)
	[0.96]	[0.154]	[0.002]	[0.951]	[0.072]	[0.002]
District FE	N	Y	Y	N	Y	Y
Subdistrict FE	N	N	Y	N	N	Y
R-squared	0.1335	0.4368	0.6728	0.0891	0.4637	0.6924
	Panel B: Outgoing Link Deletion Algorithm					
Income Correlation	0.296	0.109	-0.001	0.205	0.182	0.098
	(0.109)	(0.061)	(0.064)	(0.088)	(0.069)	(0.063)
	[0.004]	[0.12]	[0.996]	[0.03]	[0.012]	[0.13]
Income Variability	-0.061	0.015	-0.281	-0.054	-0.054	-0.304
	(0.187)	(0.239)	(0.094)	(0.206)	(0.224)	0.101
	[0.82]	[0.928]	[0.002]	[0.774]	[0.794]	[0.002]
District FE	N	Y	Y	N	Y	Y
Subdistrict FE	N	N	Y	N	N	Y
R-squared	0.1066	0.5701	0.7458	0.071	0.5897	0.7525

Notes: Outcome variable is the average Myerson distance, computed by the approximation algorithms as indicated. Panel A uses the incoming link deletion algorithm and Panel B uses the outgoing link deletion algorithm. Outcome variables and regressors scaled by their standard deviations. Columns (1-3) use Measure I of the caste-income correlation, whereas columns (4-6) use Measure II. Specifications include control for caste composition: $p(1-p)$, where p is share in high caste. Wild clustered bootstrap standard errors are presented in (.), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild

Table 3. Myerson distance vs. income variability and correlation, with village FE

	Panel A: Incoming Link Deletion Algorithm		*Panel B: Outgoing Link Deletion Algorithm*	
	Measure I	Measure II	Measure I	Measure II
	(1)	(2)	(3)	(4)
Income Correlation x **1**{Financial network}	0.242	0.177	0.225	0.289
	(0.104)	(0.115)	(0.084)	(0.115)
	[0.012]	[0.188]	[0.010]	[0.016]
Income Variability x **1**{Financial network}	-0.070	-0.057	-0.151	-0.131
	(0.107)	(0.115)	(0.178)	(0.171)
	[0.564]	[0.712]	[0.446]	[0.502]
R-squared	0.8353	0.8297	0.7362	0.7452

Notes: Outcome variable is the average Myerson distance of the graph of a given type (financial or social), computed using the specified algorithm (incoming or outgoing link deletion). Outcome variable, income correlation (either Measure I or Measure II), and income variability all scaled by their standard deviations. With two observations per village, one for each network type, all regressions include village fixed effects (and therefore subsume district and subdistrict fixed effects as well). Specifications include control for caste composition: $p(1-p)$, where p is share in high caste, interacted with network type. Wild clustered bootstrap standard errors are presented in (.), using Rademacher weights. The cluster is at the subdistrict level, of which there are 12. 1000 samples are used per bootstrap. p-values from the Wild clustered bootstrap t are presented in [.].

Table 4. Association between average within-caste centrality of nodes with cross-caste links and measures of rainfall variability and within-caste income correlation

	Incoming Link Deletion Algorithm		*Outgoing Link Deletion Algorithm*	
	Measure I	Measure II	Measure I	Measure II
	(1)	(2)	(3)	(4)
	Panel A: Using just the Financial network			
Income Correlation	-0.231	-0.196	-0.131	-0.101
	(0.098)	(0.142)	(0.061)	(0.097)
	[0.018]	[0.157]	[0.032]	[0.28]
Income Variability	0.009	-0.019	0.086	0.093
	(0.183)	(0.157)	(0.185)	(0.191)
	[0.961]	[0.918]	[0.636]	[0.609]
R-squared	0.2265	0.1992	0.127	0.111
	Panel B: Difference in Differences using the Financial and Social networks			
Income Correlation x **1**{Financial network}	-0.191	-0.228	-0.121	-0.263
	(0.088)	(0.150)	(0.095)	(0.118)
	[0.030]	[0.12]	[0.200]	[0.026]
Income Variability x **1**{Financial network}	0.056	0.060	0.109	0.104
	(0.162)	(0.166)	(0.188)	(0.177)
	[0.727]	[0.717]	[0.564]	[0.556]
R-squared	0.7956	0.7954	0.765	0.772

Notes: Outcome variable is the average within-caste average Myerson centrality, computed using our approximation algorithm (incoming link deletion or outgoing link deletion), of those with a cross-caste link. Income correlation given by Measure I or Measure II as indicated. Outcome variables and regressors are scaled by their standard deviations. Panel A's outcome variable uses only the financial network. Panel B's outcome variable is computed using the financial or social network (two observations per village for the difference-in-difference). All Panel B specifications use village fixed effects, and both panels include controls for caste composition and size (interacted with network type in Panel B). Wild clustered bootstrap standard errors are presented in (.), using Rademacher weights. The cluster is at the subdistrict level in both panels, of which there are 12. 1000 samples are used per bootstrap. *p*-values from the Wild clustered bootstrap *t* are presented in [.].

Supplementary Appendix: For Online Publication Only

Appendix A. 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 stream c_i instead of another stochastic consumption stream 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 for taking the stochastic consumption stream c_i' instead of c_i is independent of $\hat{K}$. As a villager's 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 characterize the set of Pareto efficient risk sharing agreements. Borch (1962) and Wilson (1968) 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

$$\begin{aligned}
\frac{e^{-\lambda c_i(\omega)}}{e^{-\lambda c_j(\omega)}} &= \alpha_{ij} \\
c_i(\omega)-c_j(\omega) &= -\frac{\ln(\alpha_{ij})}{\lambda} \\
\mathbf{E}[c_i(\omega)]-\mathbf{E}[c_j(\omega)] &= -\frac{\ln(\alpha_{ij})}{\lambda} \\
c_i(\omega)-c_j(\omega) &= \mathbf{E}[c_i(\omega)]-\mathbf{E}[c_j(\omega)] \qquad (6)
\end{aligned}$$

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

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

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 that,

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

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 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 agreement is Pareto efficient if and only if the sum of the consumption variances for all path-connected villagers 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 $\tau_i = \mathbf{E}[c_i(\omega)] - \mathbf{E}[\sum_{k\in\mathbf{S}} y_k(\omega)]$. □

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

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

Some additional notation will be helpful. Suppose agents arrive in an order chosen uniformly at random. The random variable $\widehat{\mathbf{S}}_i \subseteq \mathbf{N}$ identifies the set of agents, including i, who arrive weakly before i. For each arrival order, we then have an associate network $L(\widehat{\mathbf{S}}_i)$ that describes the network formed upon i's arrival. We let $\mathcal{L}(i, L)$ be the set of such networks generated by considering all possible arrival orders. Finally, let $q(i, j, L)$ be the probability that i and j are path connected on a network $L(\widehat{\mathbf{S}}_i)$ selected uniformly at random from $\mathcal{L}(i, L)$.

The certainty-equivalent value of the reduction in variance due to a link l_{ij} in a graph $L(\widehat{\mathbf{S}}_i)$ is V if the link is essential and 0 otherwise. The change in i's Myserson Value, $MV_i(L) - MV_i(L/l_{ij})$, is then just the product of the probability $1-1/2-q(i,j,L)$ and the value V. For the link l_{ij} on to be essential on a network $L(\widehat{\mathbf{S}}_i)$, j must have already arrived which occurs with probability $1/2$, and there cannot be another path from i to j. As such an alternative path exists with probability $q(i,j,L)$, and as there can only be a path from i to j if j has been added, 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 $\mathbf{P}(i,j,L') \neq \emptyset$, because a non-minimal path exists only if a minimal path exists. We therefore need to find the probability that at least one minimal path $P_k(i,j,L) \in \mathbf{P}(i,j,L)$ exists when i arrives. Let $\Pr(P_k(i,j,L))$ be the probability of the event that i is the last agent to arrive from all those involved in the path $P_k(i,j,L)$, for a random arrival order drawn from $\mathcal{L}(i,L)$. This is a necessary and sufficient condition for the path $P_k(i,j,L)$ to exist when i 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 upon i's arrival in a random draw from $\mathcal{L}(i,L)$, i.e., $\Pr\left(\cup_{P_k(i,j,L) \in \mathbf{P}(i,j,L)} P_k(i,j,L)\right)$. As these paths are not disjoint, the inclusion–exclusion principle needs to be applied. Doing so results in the formula shown.[42] □

Proof of Proposition 16. Denote the set of all possible arrival orders for the set of agents N, by $\mathcal{A}(N)$. Order this set of $|N|!$ arrival orders in any way, denoting the kth arrival order by $\widehat{A}_k \in \mathcal{A}(N)$. We will then construct an alternative ordering, in which we denote the kth arrival order by $\widetilde{A}_k \in \mathcal{A}(N)$, such that for arrival order $\widetilde{A}_k$,

(i) i arrives at the same time as agent i' does for the arrival order $\widehat{A}_k$;
(ii) when i arrives he connects to exactly the same set of agents from $\mathbf{N} \setminus \mathbf{S}_0$ that i' connects to upon his arrival for the arrival order $\widehat{A}_k$;
(iii) when i arrives he connects to weakly more agents from $\mathbf{S}_0$ that i' connects to upon his arrival for the arrival order $\widehat{A}_k$.

[42] 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'' different nodes that appear across both paths, so $K'' = |P_k(i,j,L) \cup P_{k'}(i,j,L)|$, then the probability that both paths exists when i arrives is $(1/K'')$. So, the probability that either path k or k' is present is the probability path k exists when i arrives plus the probability path k' exists when i arrives less the probability that both of them exists when i arrives: $(1/K) + (1/K') - (1/K'')$.

Equation 4 shows that the risk reduction, and hence the marginal contribution made by an agent $k \in \mathbf{S}_0$ from providing 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 construct the alternative ordering of the set $\mathcal{A}(N)$ as claimed we will directly adjust individual arrival orders, but in a way that preserves the set $\mathcal{A}(N)$. First, for each arrival order, we switch the arrival positions of i' and i. This alone is enough to ensure that conditions (i) and (ii) are satisfied. There are $|\mathbf{S}_0|!$ possible arrival orders for the set of agents $\mathbf{S}_0$. Ignoring for now the other agents, we label these arrival orders lexicographically. First we order them, in ascending order, by when i arrives. Next, we order them in ascending order by the number of agents i is connected to upon his arrival. Breaking remaining ties in any way, we have labels $1_i, 2_i, \ldots, |\mathbf{S}_0|!_i$. We then let every element of $\mathcal{A}(N)$ inherit these labels, so that two arrival orders receive the same label if and only if the agents $\mathbf{S}_0$ arrive in the same order. We now construct a second set of labels by doing the same exercise for i', and denote these labels by $1_{i'}, 2_{i'}, \ldots, |\mathbf{S}_0|!_{i'}$. We are now ready to make our final adjustment to the arrival orders. For each original arrival order $\widehat{A}_k$ we find the associated (second) label. Suppose this is $x_{i'}$. We then take the current kth arrival order (given the previous adjustment), and reorder (only) the agents in $\mathbf{S}_0$, so that the newly constructed arrival order now has (first) label x_i. Because of the lexicographic construction of the labels, the arrival position of agent i will not change as a result of this reordering of the arrival positions of agents in $\mathbf{S}_0$, so conditions (i) and (ii) are still satisfied. In addition, condition (iii) will now be satisfied from the definition of i being more central than i'. The only remaining thing to verify is that the set of arrival orders we are considering has not changed (i.e. that we have, as claimed, constructed an alternative ordering of the set $\mathcal{A}(N)$) and this also holds by construction. $\square$

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

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

Multiplying by $\lambda/2$ to get the certainty-equivalent value of the variance reduction completes the proof. $\square$

Proof of Proposition 20. We will say that agent k is a distance-t neighbor of i if the shortest path from i to k take exactly t steps (and contain $t+1$ agents, including i and k).

Consider the implementation of the Incoming Link Deletion algorithm to find $\widehat{q}_{ij}$. We begin by calculating $x^{1,i} = \mathbf{A}e^i$, where e^i is the ith basis vector. This identifies all agents connected to i. We then set all entries in the ith row from the adjacency matrix $\mathbf{A}$ to 0 and call this new matrix $\mathbf{A}_2$. This deletes the inward links to$i$ 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 rows 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 $t < l$, $x_j^{l,i} = 0$; for $t = l$, $x_j^{l,i} = 1$; and for all $t > l$, $x_j^{t,i} = 0$. Deleting incoming links ensures for all $t > l+1$, $x_j^{t,i} = 0$. As L is a tree there, there is no path of length $l+1$ to j and so $x_j^{l+1,i} = 0$. The algorithm therefore finds the unique minimal path from any i to any j and records its length; If the unique path from i to j has length l, $\widehat{q}_{ij} = 1/l$. From equation 2 it is also easily checked that $q_{ij} = 1/l$. □

Appendix B. Income Correlation from Occupation Correlations

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

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

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

(1) $\widehat{\rho_w - \rho_a}^{I} = \sum_g \phi_g \Pr(o_{i,g} = o_{j,g}) - \Pr(o_{i,A} = o_{j,B})$
(2) $\widehat{\rho_w - \rho_a}^{II} = \sum_{k=1}^{K} \mathrm{corr}\left(Caste, Occupation_k\right) \Pr\left(Occupation_k\right)$.

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

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

In what follows, assume the following:

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

where ϵ_o is a mean-0, variance-σ_o^2 iid shock that hits each occupation, u_i is an iid shock with mean zero and variance σ_u^2 that hits each individual, and $\mu_{g,o}$ is a (caste–occupation)-specific mean.

By the law of total covariance, we have

$$\begin{aligned}\operatorname{cov}(y_{i,A}, y_{j,B}) = &\sum_{o\in O}\sum_{o'\in O}\operatorname{cov}(y_{i,A}, y_{j,B}|o_{i,A}, o_{j,B})\Pr(o_{i,A} = o, o_{j,B} = o')\\ &+ \operatorname{cov}(\mathrm{E}[y_{i,A}|o_{i,A}], \mathrm{E}[y_{j,B}|o_{j,B}]),\end{aligned}$$

which we will use in our computations below.

B.1. **Identical means and variances.** We begin with the simple case where all means and variances are identical. This justifies the use of $\widehat{\rho_w - \rho_a}^I$.

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

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

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

Proof. It is immediately clear that

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

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

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

Also, note that

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

implying

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

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

$$\begin{aligned}\operatorname{corr}(y_{i,g}, y_{j,g'}) &= \frac{\operatorname{cov}(y_{i,g}, y_{j,g'})}{\sqrt{(\sigma^2+\sigma_u^2)(\sigma^2+\sigma_u^2)}}\\ &= \Pr(o_{i,g} = o_{j,g'})\frac{\sigma^2}{\sigma^2+\sigma_u^2}.\end{aligned}$$

Weighting by population share, we have

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

which completes the proof. □

B.2. **Differing average incomes by caste and occupation.**

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

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

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

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

B.3. **Differing variances by occupation.** Suppose now that

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

where ϵ_o is a mean-0, variance-σ_o^2 iid shock that hits each occupation. It follows that

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

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

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

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

Therefore, we have

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

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

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

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

$$\rho_a = \frac{\sum_{o \in O} \mathrm{Pr}(o|A)\mathrm{Pr}(o|B)\sigma_o^2}{\left(\sum_{o \in O} \mathrm{Pr}(o|A)\sigma_o\right)\left(\sum_{o' \in O} \mathrm{Pr}(o'|B)\sigma_{o'}\right)}$$

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

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

In sum,

$$\rho_w = \mathrm{Pr}(A)\frac{\sum_{o \in O} \mathrm{Pr}(o|A)^2\sigma_o^2}{\left(\sum_{o \in O} \mathrm{Pr}(o|A)\sigma_o\right)^2} + \mathrm{Pr}(B)\frac{\sum_{o \in O} \mathrm{Pr}(o|B)^2\sigma_o^2}{\left(\sum_{o \in O} \mathrm{Pr}(o|B)\sigma_o\right)^2}.$$

This implies that

$$\begin{aligned}\rho_w - \rho_a &= \Pr(A)\frac{\sum_{o\in O}\Pr(o|A)^2\sigma_o^2}{\left(\sum_{o\in O}\Pr(o|A)\sigma_o\right)^2} + \Pr(B)\frac{\sum_{o\in O}\Pr(o|B)^2\sigma_o^2}{\left(\sum_{o\in O}\Pr(o|B)\sigma_o\right)^2}\\ &\quad - \frac{\sum_{o\in O}\Pr(o|A)\Pr(o|B)\sigma_o^2}{\left(\sum_{o\in O}\Pr(o|A)\sigma_o\right)\left(\sum_{o'\in O}\Pr(o'|B)\sigma_{o'}\right)}.\end{aligned}$$

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