



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Jiang, B;Sung Kim, J;Li, C;Yang, O

Title:

Social Network Structure and Risk Sharing in Villages

Date:

2018-07

Citation:

Jiang, B., Sung Kim, J., Li, C. & Yang, O. (2018). Social Network Structure and Risk Sharing in Villages. *B.E. Journal of Economic Analysis and Policy*, 18 (3), pp.1-8. <https://doi.org/10.1515/bejeap-2017-0263>.

Persistent Link:

<https://hdl.handle.net/11343/220356>

Social Network Structure and Risk Sharing in Villages

Bin Jiang¹, Jun Sung Kim¹, Chuhui Li¹, and Ou Yang ^{*2}

¹Econometrics and Business Statistics, Monash University

²Melbourne Institute of Applied Economic and Social Research, The University of Melbourne

Abstract

This paper studies how the structure of friendship networks affects risk sharing in villages. Using techniques for partially identified econometric models, we construct a sharp bound on the true risk-sharing rate, which takes into account nomination errors in survey responses, and implement interval estimation. We show that the diameter of a network has a negative and significant impact on risk sharing. Our result implies that policymakers can effectively improve risk sharing between households by adopting policies that increase the network connectivity of individuals in the periphery of the social network.

JEL: O15, D85

Keywords: Risk sharing; Social networks; Diameter; Nomination errors

*Corresponding author: Ou Yang, Melbourne Institute of Applied Economic and Social Research, Faculty of Business and Economics, 111 Barry Street, The University of Melbourne, Victoria 3010, Australia. E-mail: ou.yang@unimelb.edu.au

1 Introduction

In times of trouble, individuals rely on friends and family. Such mutual assistance is particularly important in developing areas, where formal risk-sharing devices are scarce. Although full insurance of risks without a formal risk-sharing device is impossible (Townsend, 1994), when underlying social networks are incomplete (i.e., not all individuals are linked with one another), certain network structures are still related to better risk sharing because they may be advantageous for information processing. This paper investigates how the structures of social networks affect risk sharing in villages.

The role of social networks in risk sharing has been widely studied in the literature (see Fafchamps, 2011, for a survey). Recently, Bloch, Genicot, and Ray (2008) investigate the structure of self-enforcing insurance networks. Ambrus, Mobius, and Szeidl (2014) further study the question how local obligations and transfers aggregate up to good global risk-sharing, by developing a model in which connections between individuals serve as social collateral to enforce informal insurance payments. To our knowledge, however, only a few papers (e.g., Jackson, Rodriguez-Barraquer, and Tan, 2012) investigate which types of network structures are advantageous for risk sharing. Alatas, Banerjee, Chandrasekhar, Hanna, and Olken (2016) show that information processing in a community depends on its social network structure. Because risk sharing between households is based on information about neighbors' incomes and consumption, network structures are likely to affect risk sharing as well.

Using the “Social Networks and Microfinance” data from Banerjee, Chandrasekhar, Duflo, and Jackson (2013), we find that more expanded friendship networks lead to less risk sharing. An important policy implication of our results for developing areas with limited access to formal insurance devices, is that risk sharing between households can be improved through policies that increase the network connectivity of individuals in the periphery of the social network (e.g., a transportation subsidy for those who are in the periphery of the network to participate in social events).

Our empirical results are robust to measurement errors prevalent in risk-sharing data. In our data, it happens that two individuals would borrow money from each other, but neither of them would lend money to one another. To deal with such measurement errors, we employ techniques for partially identified econometric models and construct a bound for the true risk-sharing rate. We then apply an estimation method for an interval outcome variable (Beresteanu and Molinari, 2008). Comparison of results between estimations with and without taking into account

measurement errors demonstrates the advantage of our estimation approach that allows for an interval outcome variable.

2 Data

We use the “Social Networks and Microfinance” dataset from Banerjee et al. (2013), which contains individuals’ characteristics and social connections from 75 villages in rural India. We use the friendship (i.e., close non-relative) network and two types of risk-sharing networks: the Money network, which includes the “borrow money from” and “lend money to” relations, and the Rice–Kerosene network, which includes the “borrow rice/kerosene from” and “lend rice/kerosene to” relations.

We consider the degree distribution, the first (largest) eigenvalue, the diameter, and the clustering coefficient to represent the network structure (Jackson, Rogers, and Zenou, 2017). The degree distribution is the most natural network characteristic, in which the degree of a node is the number of links each individual in the network has to other agents. As is standard in describing distributions, we use the mean and the coefficient of variation, or equivalently, the relative standard deviation of the degree distribution to capture the village composition of individuals by popularity.¹ In particular, the mean degree measures the density of links in society and captures the fraction of possible links between all pairs of nodes that are present in a network. Higher moments of the distribution such as the variance or the coefficient of variation may have significant consequences for behavior as well. As the coefficient of variation rises, the network resembles a hub-and-spoke structure. In contrast, networks with a low coefficient of variation are closer to being regular, where individuals have similar degrees. Such variations in a network structure may have an effect on risk sharing.

The first eigenvalue of the network adjacency matrix has become a key quantity important for the study of a variety of network dynamic processes.² Let $\lambda_1(G)$ denote the first eigenvalue of the adjacency matrix G . Bollobás, Borgs, Chayes, and Riordan (2010) show that if the information transmission probability through a link is greater than or equal to $1/\lambda_1(G)$, then almost all nodes become informed,

¹We use the coefficient of variation instead of the standard deviation of the degree distribution because the standard deviation is affected by the level and therefore not a useful measure by itself. We thank a referee for pointing this out.

²An adjacency matrix is a square matrix such that its (i, j) element is one when i and j are linked, and zero otherwise.

and therefore, $\lambda_1(G)$ can be deemed positively related to information aggregation within the network (Alatas et al., 2016).

The diameter of a network measures the size of the network as determined by the longest of all the path lengths of pairs. It represents the linear size of a network, and measures the degree of a network's expansion. Its value is negatively correlated with the speed of information diffusion through a network; that is, if the diameter of a network is long, then information diffusion will be slow.

Clustering is essentially a measure of the local correlation or dependence among the locations of links. It captures to what extent the presence of a link is correlated with the presence of proximate links. The role of clustering in enforcing social norms is extensively discussed in Coleman (1988). As one of the commonly used statistics in the literature to capture this notion of clustering, the total clustering coefficient C_n of network n measures how likely one individual's two friends are connected. Hence, it is directly related to the number of friendship triangles in a network.³ The total clustering coefficient is defined as

$$C_n = \frac{3 \times \text{number of friendship triangles in network } n}{\text{number of connected triplets of nodes in network } n}. \quad (1)$$

Table 1 shows the descriptive statistics of the aforementioned five network characteristics.

Table 1: Friendship network characteristics

	Mean	SD	Minimum	Maximum
Average degree	6.116	0.521	5.195	7.429
CV of degree distribution	0.514	0.078	0.403	0.796
Diameter	12.480	1.528	9	16
First eigenvalue	14.488	4.925	7.950	30.322
Clustering coefficient	0.371	0.077	0.215	0.599

(1) CV: coefficient of variation

(2) Total: 75 villages

³An alternative measure of clustering is the average clustering coefficient in which one considers the percentage of friendship triangles out of potential triples on a node-by-node basis and then averages across nodes. Our estimation results are robust to those two different measures of clustering.

3 Empirical Framework

We let RS_n^* be the true risk-sharing rate in village n :

$$RS_n^* = \frac{\text{Number of pairs shared risks in village } n}{\text{Total number of pairs in village } n}. \quad (2)$$

We attempt to regress the true risk-sharing rate RS_n^* on village-level socio-economic characteristics x_n and on network characteristics z_n as follows:

$$RS_n^* = x_n' \beta_1 + z_n' \beta_2 + \varepsilon_n, \quad (3)$$

where ε_n is the network-level unobserved variable.

In this paper, we observe nominated risk-sharing relations rather than realized ones. Specifically, our data are based on the following survey questions: “*Who would you borrow money from?*” and “*Who would you lend money to?*” There exists a discrepancy that individual A would borrow money from individual B, but B would not nominate A to be a risk-sharing partner to whom he or she would lend money. Due to this nomination error, RS_n^* is unobserved. Instead, we observe the nominated risk-sharing rate RS_n .

To deal with the nomination error, we compute a probability bound for RS_n^* . Note that there are two types of misnomination: nominating the wrong partner or non-nominating an existing partner. We assume that the two types of misnomination occur with an equal probability, namely, the equal probabilities of indirect misclassification (EPIM), introduced by Nicoletti, Peracchi, and Foliano (2011). We let A_n^* and A_n be the true and nominated sets of risk-sharing relations in village n , respectively. Further, we define $ij \in A_n$ to denote the nominated risk-sharing relation between i and j . The EPIM assumption is then written as

$$\Pr(ij \notin A_n^* | ij \in A_n) = \Pr(ij \in A_n^* | ij \notin A_n). \quad (4)$$

By the law of total probability, the true risk-sharing rate for village n is:

$$\begin{aligned} RS_n^* &= \Pr(ij \in A_n^* | ij \in A_n) \Pr(ij \in A_n) + \Pr(ij \in A_n^* | ij \notin A_n) \Pr(ij \notin A_n) \\ &= (1 - \Pr(ij \notin A_n^* | ij \in A_n)) \Pr(ij \in A_n) + \Pr(ij \in A_n^* | ij \notin A_n) (1 - \Pr(ij \in A_n)) \\ &= (1 - \Pr(ij \notin A_n^* | ij \in A_n) - \Pr(ij \in A_n^* | ij \notin A_n)) RS_n + \Pr(ij \in A_n^* | ij \notin A_n) \end{aligned} \quad (5)$$

For each village, we compute the misnomination rate λ_n as the proportion of pairs that are nominated in only one of the borrowing or lending networks.⁴ For

⁴We implement the procedure separately for each type of risk sharing.

example, one (or both) of the two individuals in a pair wants to borrow money from the other, but neither of them wants to lend money to the other. We let λ_n be the maximum value for the probability of indirect misclassification. The true risk-sharing rate then lies between RS_n and $(1 - 2\lambda_n)RS_n + \lambda_n$.⁵

For each village, we obtain a bound $[RS_{L,n}, RS_{U,n}] = [RS_n, (1 - 2\lambda_n)RS_n + \lambda_n]$, which covers RS_n^* with probability one. With this interval-valued outcome variable $[RS_{L,n}, RS_{U,n}]$, there are many parameter vectors, each of which equally minimizes the sum of squared residuals; therefore, the model is set identified. Intuitively speaking, implementing a regression model of all possible combinations of values in the interval variable on the explanatory variables yields the estimated identified set. We employ an econometric approach for estimation and inference with interval outcome data, introduced by Beresteanu and Molinari (2008). This approach targets the entire identification set of population parameters for inference, and provides asymptotically exact size critical values with which to test hypotheses and construct confidence intervals. It readily applies to the model with interval data such as ours.

4 Results

Table 2 reports the mean, the standard deviation and the length of the bound for RS_n^* across 75 villages. For monetary risk sharing (Money), the length of the bound is 0.09 percentage points on average, with a range between 0.025 and 0.234 percentage points. Given that risk-sharing relations are sparse (e.g., the average nominated risk-sharing rate is 0.769%), these bounds are non-negligible.⁶

⁵Note that, if $RS_n \leq 0.5$, as in our data, then $(1 - 2\lambda_n)RS_n + \lambda_n \geq RS_n$.

⁶The results are similar for the Rice-Kerosene network.

Table 2: Bounds for the risk-sharing rate

	Money			
	Mean	SD	Min	Max
Misnomination rate	0.090%	0.045%	0.025%	0.242%
RS_U	0.857%	0.283%	0.426%	1.921%
RS_L	0.769%	0.249%	0.398%	1.696%
Interval length (percentage points)	0.088	0.044	0.025	0.234
	Rice–Kerosene			
	Mean	SD	Min	Max
Misnomination rate	0.071%	0.044%	0.017%	0.0023%
RS_U	0.794%	0.267%	0.412%	1.751%
RS_L	0.724%	0.240%	0.384%	1.620%
Interval length (percentage points)	0.070	0.043	0.016	0.224

⁽¹⁾ SD: standard deviation

⁽²⁾ Total: 75 villages

We run both ordinary least squares (OLS) estimation (with the assumption that the nominated risk-sharing rate was true) and interval estimation for the two types of risk sharing to ensure robustness. Results are similar for both types of risk sharing, and therefore, we focus on monetary risk sharing. We only report the estimation results for specifications that include village-level socio-demographic control variables because the results are almost identical without those control variables.

The OLS and interval estimation results are respectively shown in Tables 3 and 4, where Columns (1)–(5) report results from five separate regressions with network characteristics included one by one, and Column (6) presents estimates for a regression equation, where all the network variables are included at the same time. As discussed in Alatas et al. (2016), given the correlation among the network variables, the partial correlations conditional on other network characteristics are generally complicated and cannot be clearly interpreted.⁷ Therefore, our interpretations and policy implication are primarily based on the results from the one-by-one specifications.

⁷We thank a referee for pointing this out.

Table 3: Results (OLS estimation)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. The effect of network structure on risk sharing (Money)</i>						
Average degree	-0.0003 (0.0006)					-0.0002 (0.0010)
CV of the degree distribution		-0.0023 (0.0043)				0.0272*** (0.0081)
Diameter			-0.0007*** (0.0002)			-0.0008*** (0.00019)
First eigenvalue				-0.00012** (0.00006)		-0.0007*** (0.00016)
Clustering coefficient					-0.0034 (0.0038)	0.0223*** (0.0054)
R^2	0.065	0.065	0.220	0.111	0.070	0.435
Controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>Panel B. The effect of network structure on risk sharing (Rice–Kerosene)</i>						
Average degree	-0.0001 (0.0006)					0.0001 (0.0010)
CV of the degree distribution		-0.0020 (0.0041)				0.0242*** (0.0081)
Diameter			-0.0006*** (0.00016)			-0.0008*** (0.0002)
First eigenvalue				-0.0001* (0.00006)		-0.0007*** (0.0002)
Clustering coefficient					-0.0022 (0.0037)	0.0218*** (0.0054)
R^2	0.060	0.064	0.227	0.101	0.064	0.436
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	75	75	75	75	75	75

(1) CV: coefficient of variation

(2) Controls: average age, average education, and proportions of males, general caste individuals and village natives.

(3) *** p<0.01, ** p<0.05, * p<0.1

Table 4: Results (interval estimation)

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. The effect of network structure on risk sharing (Money)</i>						
Average degree						[-0.0015, 0.0014]
CV of the degree distribution		[-0.0067, 0.0020]				[0.0168, 0.0424]**
Diameter			[-0.0010, -0.0005]***			[-0.0011, -0.0005]***
First eigenvalue				[-0.0002, -0.0001]		[-0.0010, -0.0005]***
Clustering coefficient					[-0.0099, 0.0004]	[0.0112, 0.0314]***
Controls	Yes	Yes	Yes	Yes	Yes	Yes
<i>Panel B. The effect of network structure on risk sharing (Rice-Kerosene)</i>						
Average degree						[-0.0007, 0.0016]
CV of the degree distribution		[-0.0055, 0.0016]				[0.0154, 0.0364]**
Diameter			[-0.0009, -0.0005]***			[-0.0010, -0.0005]***
First eigenvalue				[-0.0002, -0.0001]		[-0.0009, -0.0005]***
Clustering coefficient					[-0.0079, 0.0009]	[0.0119, 0.0281]***
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	75	75	75	75	75	75

(1) Each interval shows the projection of the identified set of all the parameters to the one-dimensional space of the corresponding coefficient. Significance is based on the $1 - \alpha\%$ confidence interval constructed as the projection of the $1 - \alpha\%$ confidence set of parameters. For example, the 95% confidence set for each parameter is larger than the identified set shown above.

(2) CV: coefficient of variation

(3) Controls: average age, average education, and proportions of males, general caste individuals and village natives.

(4) *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

When the network variables are included one by one, the diameter has a negative and significant effect on risk sharing, as shown in both OLS and interval estimations. If a network diameter increases by one degree, the risk-sharing rate in Money will decrease by between 0.05 and 0.09 percentage points. Given that the average risk-sharing rate is 0.769%, such an increase in the diameter will result in at least a 6.5% decrease in risk sharing, which is economically significant.⁸ The diameter of a network is closely related to the expansiveness of the social network and depends on the presence of individuals whose network connectivity is substantially lower than others due to high costs of attachment to socially or geographically distant nodes (Jackson and Rogers, 2005). Hence, the negative coefficient on the diameter indicates that less risk sharing occurs if relatively isolated individuals exist in a village network.

The negative sign for the first eigenvalue implies that when village information is more precisely aggregated, risk sharing appears to decrease, which seems counterintuitive because asymmetric information on the income of others is related to low risk sharing in the literature (e.g., Thomas and Worrall, 1990). However, as shown in Column (2) in Table 4, once we take into consideration the nomination errors in the regression, the first eigenvalue becomes insignificant. This result implies that ignoring the nomination errors may mislead the effect of network variables on risk sharing, and demonstrates the advantage of the interval estimation approach when measurement errors are present in risk-sharing data.

The average degree does not have a significant effect on risk sharing. Although more connections typically help better communication, having a higher average degree may not guarantee better information aggregation (Alatas et al., 2016). Similarly, our finding indicates that the average degree is not strongly associated with risk sharing. The coefficient of variation is also insignificant, implying that regular networks in which individuals have similar degrees and networks that resemble a hub-and-spoke structure have no statistically different risk sharing rates. The clustering coefficient also has no significant effect on risk sharing.

Including all network variables in the regression model as shown in Column (6) of Tables 3 and 4, we once again find a negative effect of the diameter on risk sharing. In contrast, the coefficient of variation of the degree distribution and the clustering coefficient appear to be positively correlated with risk sharing. The reason we obtain different results here from those shown in the one-by-one regressions might be explained as follows. Holding all the other network characteristics constant, raising clustering leads to more risk sharing. However, an increase in

⁸Note that the difference in diameter between the longest and shortest networks is 7 degrees.

clustering of a network may not give a positive effect if such an increase changes other patterns of the network.

Our analysis with the one-by-one specifications has an important policy implication for developing areas, where formal microfinance devices are rare. Even without access to formal insurance, policymakers can improve risk sharing between households by adopting policies that increase the network connectivity of individuals in the periphery of the social network. Such policies include: supporting village leaders to regularly meet rather isolated individuals and subsidizing transportation costs to individuals in the periphery of the network to encourage their participation in social events.

It should be cautioned that the interpretation of our findings is constrained by the survey design. There are two potential caveats of our study to consider. First, the survey only covered roughly half of the population in each village, that is, the other half do not have a chance to name who they would borrow money from or lend money to. Hence, there is a potential issue due to missing information. Second, there may be a trade-off between the quantity and quality of risk-sharing friends, so having more risk-sharing friends does not necessarily mean that there is a higher degree of insurance against idiosyncratic risk. However, in the survey, we were not able to find questions which relate to and can thus be used to control for the quality of risk-sharing friends. Further research may be carried out to extend this study by expanding the survey to cover the whole population and adding questions which can be used to measure the quality of risk-sharing friends to the survey questionnaire.

5 Conclusion

This paper examines how the structure of friendship networks affects risk sharing in villages. To yield results that are robust to nomination errors prevalent in risk-sharing data, the techniques for partially identified econometric models are employed, and the interval estimation is carried out for the two types of risk sharing, the Money network and the Rice-Kerosene network. Our results indicate that more expanded friendship networks lead to less risk sharing. Accordingly, policymakers can effectively increase risk sharing between households by adopting a policy that improves the network connectivity of individuals in the periphery of the social network in developing areas even when no formal insurance devices are available.

References

- Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2016): “Network structure and the aggregation of information: Theory and evidence from indonesia,” *American Economic Review*, 106, 1663–1704.
- Ambrus, A., M. Mobius, and A. Szeidl (2014): “Consumption risk-sharing in social networks,” *American Economic Review*, 104, 149–82.
- Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): “The diffusion of microfinance,” *Science*, 341, 363–370.
- Beresteanu, A. and F. Molinari (2008): “Asymptotic properties for a class of partially identified models,” *Econometrica*, 76, 763–814.
- Bloch, F., G. Genicot, and D. Ray (2008): “Informal insurance in social networks,” *Journal of Economic Theory*, 36–58.
- Bollobás, B., C. Borgs, J. Chayes, and O. Riordan (2010): “Percolation on dense graph sequences,” *The Annals of Probability*, 38, 150–183.
- Coleman, J. S. (1988): “Social capital in the creation of human capital,” *American Journal of Sociology*, 94, S95–S120.
- Fafchamps, M. (2011): “Risk sharing between households,” *Handbook of Social Economics*, 1, 1255–1279.
- Jackson, M. O., T. Rodriguez-Barraquer, and X. Tan (2012): “Social capital and social quilts: Network patterns of favor exchange,” *The American Economic Review*, 102, 1857–1897.
- Jackson, M. O. and B. W. Rogers (2005): “The economics of small worlds,” *Journal of the European Economic Association*, 3, 617–627.
- Jackson, M. O., B. W. Rogers, and Y. Zenou (2017): “The economic consequences of social-network structure,” *Journal of Economic Literature*, 55, 49–95.
- Nicoletti, C., F. Peracchi, and F. Foliano (2011): “Estimating income poverty in the presence of missing data and measurement error,” *Journal of Business & Economic Statistics*, 29, 61–72.

Thomas, J. and T. Worrall (1990): "Income fluctuation and asymmetric information: An example of a repeated principal-agent problem," *Journal of Economic Theory*, 51, 367–390.

Townsend, R. (1994): "Risk and insurance in village India," *Econometrica*, 62, 539–91.