Informal Risk Sharing with Local Information

Attila Ambrus[*], Wayne Y. Gao[†], Pau Milán[‡]

October 31, 2016

Abstract

This paper considers the effect of local information constraints in risk-sharing networks. We assume individuals only observe the endowment realizations of their neighbors, and risk-sharing arrangements between two individuals can only depend on commonly observed information. We derive necessary and sufficient conditions for Pareto efficiency subject to these constraints. We provide an explicit characterization of Pareto efficient arrangements under CARA utilities and normally distributed endowments. With independent endowments, local equal sharing rules are shown to be optimal. For correlated endowments, the optimal sharing rule is characterized in closed-form as a function of a network measure of centrality. Contrary to other models of informal insurance in networks, more central individuals are likely to become quasi insurance providers to more peripheral individuals, and attain more volatile consumption. We argue that the current framework has important implications for empirical tests of risk-sharing, and that standard estimates of risk-sharing tests may be decomposed into an underlying heterogeneity of insurance opportunities that can be interpreted economically in terms of consumption volatility.

Keywords: social network, risk sharing, Pareto efficiency, local information

[*]Attila Ambrus: Department of Economics, Duke University, Durham, NC 27708, USA, attila.ambrus@duke.edu.

[†]Wayne Yuan Gao: Department of Economics, Yale University, New Haven, CT 06511, USA, wayne.gao@yale.edu.

[‡]Pau Milán: Department of Economics, Universitat Autnoma de Barcelona, Edifici B, Bellaterra 08193, Spain, pau.milan@uab.cat.

1 Introduction

Informal insurance arrangements in social networks have been shown to play an important role at smoothing consumption in a number of different contexts (Rosenzweig (1988), Deaton (1992), Paxson (1993), Udry (1994), Townsend (1994), Grimard (1997), Fafchamps and Lund (2003) and Fafchamps and Gubert (2007a)). A main finding in this literature is that informal insurance achieves imperfect consumption smoothing. There are different theoretical explanations as to why perfect risk sharing is not possible. One leading explanation is the presence of enforcement constraints: since risk-sharing arrangements are informal, they have to satisfy incentive compatibility, implying an upper bound on the amount of transfer that individuals can credibly promise to each other. This type of explanation has been explored in a social network framework by Ambrus, Mobius, and Szeidl (2014).[1]

In this paper we explore an alternative explanation featuring local information constraints: individuals can only observe endowment realizations of their direct neighbors, and insurance arrangements between two connected individuals can only be conditioned on local information, consisting of the endowment realizations of their common neighbors (including themselves).[2] In contrast, existing models of informal risk sharing in networks (Bramoullé and Kranton (2007), Ambrus, Mobius, and Szeidl (2014), Ambrus, Chandrasekhar, and Elliott (2015)) assume that any bilateral arrangement between connected individuals can be conditioned on global information, meaning the community's full set of endowment realizations.[3] We find that this explanation generates qualitatively different predictions that are empirically testable relative to models of informal insurance with enforcement constraints. Hence, our

[1]See also Karlan, Mobius, Rosenblat, and Szeidl (2009), who investigate enforcement constraints in the case of a single borrowing transaction. There is also an extensive literature on the effects of limited commitment on risk-sharing possibilities for a pair of individuals instead of general networks (Coate and Ravallion (1993), Kocherlakota (1996), Ligon (1998), Fafchamps (1999), Ligon, Thomas, and Worrall (2002), Dubois, Jullien, and Magnac (2008)).

[2]Throughout the paper we maintain the terminology "individuals", even though in many contexts the relevant unit of analysis is households.

[3]Bloch et al. (2008) consider various exogenously-specified transfer rules that do not have to depend on all endowment realizations, but a transfer between two individuals does depend on transfers to or from other individuals - that is on nonlocal information. For example, global equal sharing at a network component level is feasible in the framework of Bloch et al. (2008) , while it is not for general networks in our model. See also Bourlès, Bramoullé, and Perez-Richet (2016), where individuals are motivated to send transfers to their neighbors for explicit altruistic reasons, but like in Bloch et al. (2008) , bilateral transfers depend on transfers among other individuals.

results can help future empirical work identify which type of constraint plays the key role in keeping informal insurance arrangements from the efficient frontier.[4] Most recently, Milán (2016) performs an empirical analysis of this model's implications using data from indigenous communities in the Bolivian Amazon basin.

There is a line of theoretical literature investigating the effect of imperfect observability of incomes on informal risk sharing arrangements between two individuals in isolation: see for example Townsend (1982), Thomas and Worrall (1990), and Wang (1995). The questions investigated in this literature are fundamentally different from the ones we focus on, mainly because we are interested in questions that are inherently network related.[5]

The current framework also speaks to an ongoing debate in the development literature that emphasizes the importance of appropriately defining individuals' risk-sharing groups in empirical work. (Mazzocco and Saini (2012), Angelucci, de Giorgi, and Rasul (2015), Attanasio, Meghir, and Mommaerts (2015), Munshi and Rosenzweig (2016)). A general trend in this literature considers alternative sub-groups within communities as the relevant risk-sharing units of individuals (e.g. an individual's caste or extended family). They argue that classical empirical tests of risk sharing (Townsend (1994)) must be adapted to accommodate heterogeneity in individuals' risk sharing communities. Unfortunately, they only allow for a reduced and rigid form of heterogeneity in which group membership is mutually exclusive and groups do not interact among themselves. In this respect, we present a framework that incorporates these partition models, but also allows for much more general forms of interactions in which individuals' sharing groups can overlap in complicated ways along the network. We argue that the current framework has important implications for these new empirical tests of risk-sharing. We show that not defining appropriate local groups biases results, and that standard estimates of risk-sharing tests may be decomposed into an underlying distribution of insurance opportunities that can be interpreted economically in terms of consumption volatility.

The first part of our analysis characterizes Pareto efficient risk-sharing arrange-

[4]Empirical papers trying to distinguish among different reasons of imperfectness of informal insurance contracts include Kinnan (2011) and Karaivanov and Townsend (2014). For an empirical test between full insurance versus informational constraints, see Ligon (1998).

[5]Other differences include that our analysis is static while the above papers are inherently dynamic, and in our paper individuals perfectly observe local information (but not beyond), while in the above papers incomes are not observable even between two connected individuals.

ments under local information constraints for general (concave) utility functions and endowment distributions. In the benchmark model with global information, the necessary and sufficient condition for Pareto optimality can be derived relatively easily. This leads to a simple necessary and sufficient condition for optimality, referred to as the Borch rule (Borch (1962), Wilson (1968)), stating that the ratios of any two individuals' marginal utilities of consumptions must be equalized across states. Characterizing the set of Pareto efficient arrangements subject to local information constraints is technically more challenging, as such constraints are intertwined (different pairs of connected individuals are allowed to condition their transfers on different subvectors of the state). Nevertheless, we can generalize the Borch rule to this setting. In particular, a necessary and sufficient condition for Pareto optimality of a risk-sharing arrangement with local information equates the ratios of *expected* marginal utilities of consumption for each linked pair, where expectations are conditional on local states (i.e. on the realizations of commonly observed endowments). As in the context of risk-sharing arrangements with global information, it can be shown that for each set of Pareto-weights, there is a unique Pareto efficient consumption plan, characterized by the expectation-form Borch rule. [6]

The generalized Borch rule can be used to verify the Pareto efficiency of consumption plans achieved by candidate transfer agreements in concrete specifications of our model. We provide this characterization for the case of CARA utilities and jointly normally distributed endowments. The characterization is particularly simple for independent endowments: each individual shares her endowment realization equally among her neighbors and herself; on top of that, the arrangement can include state independent transfers, affecting the distribution of surplus but not the aggregate welfare. This type of transfer arrangement, which we refer to as the local equal sharing rule, was considered as an ad hoc sharing rule in Gao and Moon (2016). Our result provides micro-foundations for the rule, in the context of CARA utilities and independent normally distributed endowments. The rule is particularly simple in that bilateral transfers are linear in endowment realizations and they only depend on the pair's endowment realizations, not on those of common neighbors.

For the more general case of correlated endowment realizations in the CARA-

[6] Just like in the standard setting with risk-sharing arrangements that can be conditioned on global information, for arrangements that can only be conditioned on local information it also holds that the set of Pareto efficient risk-sharing arrangements are equivalent to the set of solutions for a utilitarian planner's problem, for different weights.

normal setting, we show that efficient risk-sharing can still be achieved by transfers that are linear in endowment realizations and strictly bilateral (i.e. transfers that only depend on the endowment realizations of the pair involved). Moreover, we derive a closed form solution for these efficient transfers for any given network structure. In contrast to the local equal sharing rule that obtains in the case of independent endowments, we find that if individuals i and j are connected, increasing the exposure of i to transfers from non-common neighbors increases the share of i's endowment realization transferred to j, relative to local equal sharing, and decreases the share of j's endowment realization transferred to i. These correction terms, which are complicated functions of the network structure, take into account that more centrally located individuals are more exposed to the common shock component, and optimally correct for this discrepancy. These correction effects can be summarized in two different measures of an individual's network centrality. These measures depend on the individual's network position, as well as the correlation coefficient between endowment realizations. In particular, one centrality measure is related to the fraction of her own endowment realization an individual keeps, relative to the transfers sent to other neighbors, while the other centrality measure is related to the fraction of the neighbors' endowment the individual receives. In general, more central individuals, according to these measures, transfer a higher share of their endowment realizations to less central neighbors, and conversely take a smaller fraction of neighbors' endowment realizations.

Interestingly, since the centrality measures can be computed from observable variables, these predictions can be tested empirically. In fact, Milán (2016) tests the pairwise transfer scheme predicted by local information constraints against the observed exchanges of food between households of the Bolivian Amazon. The analysis shows that the network-based transfers under local information are good predictors of the pairwise sharing behavior of these aboriginal communities, and that local information constraints can adequately account for the departure of these communities from the efficient benchmark.

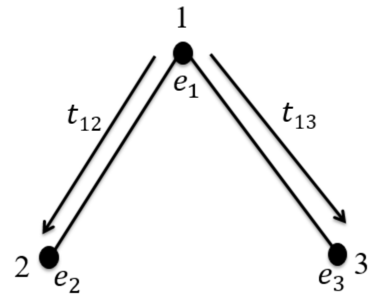
Even with the above correction terms relative to local equal sharing, more central individuals end up with a higher consumption variance because they serve as quasi insurance providers to more peripheral neighbors. For a fixed set of welfare weights, they are compensated for this service through higher state-independent transfers ("insurance premium"). This is contrary to the predictions from models with enforcement constraints, like AMS, in which more centrally connected individuals are better insured (end up with smaller consumption variance) because for typical endowment realizations they end up on larger "risk-sharing islands." While the centrality measures delivered by our model are not equivalent to more standard notions of centrality, still we show that, for typical networks, the contrast between the predictions of our model and the AMS model remains when we use degree or eigenvector centrality as a substitute. We show this through simulations, using network data from two different data sources from Indian villages.[7] With simulated endowment realizations, the AMS model produces a negative correlation between either degree or eigenvalue centrality and consumption variance, more starkly for relatively tight capacities, while the optimal risk-sharing arrangements in our model yield positive correlations.

Our model also explains why informal insurance might perform badly in one setting but quite well in another, despite similarities in network structure and in the correlation across endowments. In particular, with local information constraints, high correlation between neighboring individuals' endowments substantially hurts risk-sharing efficiency. We demonstrate this property in a circle network, under the assumption that correlations between endowments geometrically decay with distance. Keeping the same level of insurable risk under the global information risk-sharing arrangements benchmark, and focusing on the case of a large number of individuals and highly correlated endowment realizations, we show that risk-sharing arrangements with local information improve very little over autarky under the decaying correlation structure, while fairly good risk-sharing can be achieved under the symmetric correlation structure (correlation not depending on social distance). This might help explain why, even though empirical research has found that informal insurance works well in many contexts, Kazianga and Udry (2006) found a setting in which informal insurance does not seem to help, and Goldstein, de Janvry, and Sadoulet (2001) found that certain types of shocks are not well insured through informal risk sharing. Moreover, we think that a setting in which correlations across endowments decay over the network distance captures a very realistic phenomenon in risk sharing contexts, where individuals that are "socially close" will tend to engage in similar productive activities.[8]

[7]The data was provided to us by Erica Field and Rohini Pande, and by Abhijit Banerjee, Arun Chandrasekhar, Esther Duflo and Matthew Jackson.

[8]See for instance, Hooper (2011), Fafchamps and Lund (2003), etc.

Figure 1: Simple 3-Individual Network



The rest of this paper is organized as follows. In Section 2, we use a simple example to illustrate the difference between risk sharing with global versus local information. In Section 3, we derive a set of necessary and sufficient conditions for Pareto efficient transfer arrangements subject to the local information constraints, for general specifications of our model. In Section 4 we explicitly characterize Pareto efficient transfer arrangements subject to local information constraints in settings with CARA utilities and jointly normally distributed endowments, and examine properties of these arrangements. In Section 5 we outline the theory's implications for empirical tests of risk-sharing, and we show how standard estimates can be decomposed into underlying heterogeneity and interpreted economically. In Section 6 we discuss a number of extensions and generalizations. Finally, Section 7 concludes. Proofs and lemmas are provided in the Appendix A and supplementary materials are available in Appendix B.

2 Example: Three Individuals in a Line Network

2.1 Specification

Before investigating general network structures, we first consider the simplest non-trivial network, when three individuals, denoted by 1, 2 and 3, are minimally connected. In particular, individual 1 is linked with individual 2 and individual 3, but individual 2 is not linked with individual 3. Despite its simplicity, this example provides some useful insights on how local information constraints affect efficient risk-sharing arrangements.

We assume that individuals have homogeneous CARA preferences of the form

$u(x) = -e^{-rx}$. Furthermore, assume that endowments e_1, e_2, e_3 are independent and normally distributed, with mean 0 and variance σ^2. Only linked individuals may enter into risk-sharing arrangements to insure against endowment risks. Let t_{12} denote the ex-post transfer from individual 1 to individual 2, t_{13} the transfer from individual 1 to individual 3. Let x_1, x_2, x_3 denote the final consumption to individuals after the transfers are implemented, i.e., $x_1 = e_1 - t_{12} - t_{13}$, $x_2 = e_2 + t_{12}$ and $x_3 = e_3 + t_{13}$.

Below we compare Pareto efficient transfer rules in two cases: under risk-sharing arrangements between neighbors that can condition transfers on everyone's endowment realization (global information), and under risk-sharing arrangements that can only condition on endowment realizations of the two individuals forming the arrangement (local information).

2.2 Risk-sharing Arrangements with Global Information

First we consider the benchmark case when bilateral risk-sharing arrangements between neighbors can be conditioned on global information, that is on all three individuals' endowment realizations, so that t_2, t_3 can be arbitrary functions of the endowments e_1, e_2, e_3. Standard arguments (see Wilson (1968)) establish that Pareto efficient transfer rules t_{12}, t_{13} are the ones maximizing the social planner's problem:

$$\mathbb{E}\left[\sum_{i=1}^{3} \lambda_i u(x_i)\right] = \mathbb{E}\left[\lambda_1 u(e_1 - t_{12} - t_{13}) + \lambda_2 u(e_2 + t_{12}) + \lambda_3 u(e_3 + t_{13})\right], \tag{1}$$

for some $\lambda_1, \lambda_2, \lambda_3 \in [0,1]$ s.t. $\lambda_1 + \lambda_2 + \lambda_3 = 1$. By the well-known Borch rule (Borch (1962), Wilson (1968)), the necessary and sufficient conditions for optimality are:

$$\lambda_1 u'(e_1 - t_{12} - t_{13}) = \lambda_2 u'(e_2 + t_{12}) = \lambda_3 u'(e_3 + t_{13}) \;\; \forall e_1, e_2, e_3.$$

With CARA utility, this yields

$$\begin{cases} t_{12}(e_1, e_2, e_3) = & \frac{1}{3}e_1 - \frac{2}{3}e_2 + \frac{1}{3}e_3 - \frac{1}{3r}\ln\left(\lambda_2^2/\lambda_1\lambda_3\right) \\ t_{13}(e_1, e_2, e_3) = & \frac{1}{3}e_1 + \frac{1}{3}e_2 - \frac{2}{3}e_3 - \frac{1}{3r}\ln\left(\lambda_3^2/\lambda_1\lambda_2\right) \end{cases}$$

and the implied final consumption for any (e_1, e_2, e_3) are:

$$\begin{cases} x_1 = \frac{1}{3}(e_1 + e_2 + e_3) + \frac{1}{3r}\ln\left(\lambda_2\lambda_3/\lambda_1^2\right) \\ x_2 = \frac{1}{3}(e_1 + e_2 + e_3) + \frac{1}{3r}\ln\left(\lambda_1\lambda_3/\lambda_2^2\right) \\ x_3 = \frac{1}{3}(e_1 + e_2 + e_3) + \frac{1}{3r}\ln\left(\lambda_1\lambda_2/\lambda_3^2\right). \end{cases} \tag{2}$$

That is, Pareto efficient risk-sharing arrangements in every state divide total realized endowments equally among individuals, and the equal division is then corrected by state-independent transfers that achieve the welfare weights.

2.3 Risk-sharing Arrangements with Local information

Suppose now that endowment realizations are only locally observable and verifiable, so that the transfers t_{12}, t_{13} in the risk-sharing arrangements can be contingent on local information only, that is:

$$t_{12} = t_{12}(e_1, e_2),\ t_{13} = t_{13}(e_1, e_3).$$

Achieving consumption plans on the Pareto frontier, given by (2), is impossible when risk-sharing arrangements can only condition on local information. However, a necessary condition for a transfer arrangement to be socially optimal is that, for any given realization of e_1 and e_2, t_{12} should maximize $\lambda_1 u(e_1 - t_{12} - t_{13}) + \lambda_2 u(e_2 + t_{12})$, given the distribution of e_3 conditional on e_1 and e_2, and the implied distribution of consumption levels (net of t_{12}) induced by t_{13}. In short, given t_{13}, t_{12} should maximize the planner's welfare function:

$$\max_{t_{12}} \int \left[\lambda_1 u(e_1 - t_{12} - t_{13}) + \lambda_2 u(e_2 + t_{12})\right] f_{3|12}(e_3)\, de_3$$

The necessary and sufficient FOC for this maximization problem is:

$$\lambda_1 \mathbb{E}\left[u'(e_1 - t_{12} - t_{13}(e_1, e_3)) \middle| e_1, e_2\right] = \lambda_2 u'(e_2 + t_{12}). \tag{3}$$

and similarly for t_3 given t_2. Solving this system of two integral equations, we obtain the Pareto efficient transfers:

$$\begin{cases} t_{12}(e_1,e_2) = \frac{1}{3}e_1 - \frac{1}{2}e_2 - \frac{1}{24}r\sigma^2 - \frac{1}{3r}\ln\left(\lambda_1\lambda_3/\lambda_2^2\right) \\ t_{13}(e_1,e_3) = \frac{1}{3}e_1 - \frac{1}{2}e_3 - \frac{1}{24}r\sigma^2 - \frac{1}{3r}\ln\left(\lambda_1\lambda_2/\lambda_3^2\right) \end{cases}. \tag{4}$$

Notice the transfers can be decomposed into three parts. The first part, $\frac{1}{3}e_1 - \frac{1}{2}e_2$, corresponds to the "local equal sharing rule", which is the local variant of the equal sharing rule. It implies that individual i's endowment e_i is equally shared by i and i's neighbors, i.e., $t_{ij} = \frac{e_i}{d_i+1} - \frac{e_j}{d_j+1}$. The second part of the equations in (4), $-\frac{1}{24}r\sigma^2$, corresponds to a state-independent transfer that can be regarded as the "insurance premium" paid by the "net insurance purchaser" to the "net insurance provider". In this case, as the final consumption are

$$\begin{cases} x_1 = \frac{1}{3}e_1 + \frac{1}{2}e_2 + \frac{1}{2}e_3 + \frac{1}{12}r\sigma^2 + \frac{1}{3r}\ln\left(\lambda_1^2/\lambda_2\lambda_3\right), \\ x_2 = \frac{1}{3}e_1 + \frac{1}{2}e_2 - \frac{1}{24}r\sigma^2 + \frac{1}{3r}\ln\left(\lambda_2^2/\lambda_1\lambda_3\right), \\ x_3 = \frac{1}{3}e_1 + \frac{1}{2}e_3 - \frac{1}{24}r\sigma^2 + \frac{1}{3r}\ln\left(\lambda_3^2/\lambda_1\lambda_2\right), \end{cases}$$

individual 1 takes extra risk exposure $\frac{1}{3}e_1 + \frac{1}{2}e_2 + \frac{1}{2}e_3$ in comparison to individuals 2 and 3, $\frac{1}{3}e_1 + \frac{1}{2}e_2$ or $\frac{1}{3}e_1 + \frac{1}{2}e_3$. Hence, individual 1 is rewarded the certainty equivalent (CE) for her intermediary role in risk sharing. The third part of the equations in (4), $-\frac{1}{3r}\ln\left(\lambda_1\lambda_2/\lambda_3^2\right)$, redistributes wealth according to the welfare weights assigned to different individuals (it is zero when $\lambda_1 = \lambda_2 = \lambda_3$).

To evaluate the welfare loss associated with risk-sharing arrangements to be conditional only on local information, note that because social welfare is a linear, strictly decreasing function of total variances under CARA utilities, we can simply compare the total variances of final consumption. With global information, the sum of consumption variances is: $TVar_G = 3 \cdot Var\left[\frac{1}{3}(e_1 + e_2 + e_3)\right] = \sigma^2$. With bilateral risk-sharing arrangements subject to the local information constraints, the sum of consumption variances increases to: $TVar_L = Var\left[\frac{1}{3}e_1 + \frac{1}{2}e_2 + \frac{1}{2}e_3\right] + Var\left[\frac{1}{3}e_1 + \frac{1}{2}e_2\right] + Var\left[\frac{1}{3}e_1 + \frac{1}{2}e_3\right] = \frac{4}{3}\sigma^2$. Hence the welfare loss arising from local information constraints is $\frac{1}{3}\sigma^2$ in the above example.

3 General Conditions for Pareto Efficiency

Before we proceed to our main analysis, we introduce some notations. Let $N = \{1, 2, ..., n\}$ be a finite set of individuals and let G be the adjacency matrix of a network structure on N. A pair of individuals i, j are linked if $G_{ij} = 1$, and by convention, $G_{ii} = 0$. Throughout the paper we assume, without loss of generality, that G represents a connected network.[9] Define the neighborhood of i $N_i := \{j \in N : G_{ij} = 1\}$ and the extended neighborhood of i $\overline{N}_i := N_i \cup \{i\}$. The degree of individual i is defined as $d_i := \#(N_i)$, the number of individuals to whom i is linked. Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, we model each individual's endowment as a random variable e_i defined on $(\Omega, \mathscr{F}, \mathbb{P})$, and we denote the joint distribution of the vector of endowments $e \equiv (e_i)_{i \in N}$ by P. We use $\mathbb{E}[\cdot]$ to denote the expectation operator under the probability measure $\mathbb{P}$. Furthermore, we sometimes abuse notations and write $\omega \equiv e(\omega) \equiv (e_1(\omega), e_2(\omega), ..., e_n(\omega))$, i.e. we treat the state of world as interchangeable with the joint realization of endowments, taking $\mathscr{F} = \mathscr{B}(\mathbb{R}^n)$.

A central assumption in our paper is that individuals can only observe the endowment realizations of their direct neighbors. Define $N_{ij} := N_i \cap N_j$ and $\overline{N}_{ij} := \overline{N}_i \cap \overline{N}_j$. Let $I_i := (e_j)_{j \in \overline{N}_i}$ be the information vector of i. Then the common information of individuals i and j is $I_{ij} := (e_k)_{k \in \overline{N}_{ij}}$.

We assume that only linked pairs of individuals can engage in informal risk sharing directly. Risk sharing takes the form of ex ante risk-sharing arrangements between linked individuals i and j on a net transfer t_{ij} from i to j, conditional on the realization of the state. We assume that individuals can only condition the transfer on their common information. Formally, we require that $t_{ij} : \Omega \to \mathbb{R}$ be $\sigma(I_{ij})$-measurable, where $\sigma(I_{ij})$ denotes the σ-algebra induced by I_{ij}. By definition, $t_{ij}(\omega) = -t_{ji}(\omega)$ for every $\omega \in \Omega$ and linked $i, j \in N$. We refer to the collection of ex ante risk-sharing arrangements of the above form between all pairs of linked individuals as a transfer arrangement.

Let $\mathcal{T}$ denote the set of all possible transfer arrangements $t : \Omega := \mathbb{R}^n \to \mathbb{R}^{\sum_{i \in N} d_i}$ that are only contingent on commonly known information for each linked pair:

$$\mathcal{T} := \left\{ t : \Omega \to \mathbb{R}^{\sum_{i \in N} d_i} \;\middle|\; \forall i, j \text{ s.t. } G_{ij} = 1, \quad \begin{array}{c} t_{ij} \text{ is } \sigma(I_{ij})\text{-measurable} \\ \text{and } t_{ij}(\omega) + t_{ji}(\omega) = 0, \ \forall \omega. \end{array} \right\}$$

[9] Otherwise we may analyze each component separately.

Define $\mathcal{T}^* \subseteq \mathcal{T}$ by $\mathcal{T}^* := \{t \in \mathcal{T} | \langle t, t \rangle < \infty\}$, where $\langle s, t \rangle := \mathbb{E}\left[\sum_{G_{ij}=1} s_{ij}(\omega) t_{ij}(\omega)\right]$ for any $s, t \in \mathcal{T}$. It follows that $\langle \cdot, \cdot \rangle$ is an inner product and $\mathcal{T}^*$ is a well-defined Hilbert space (see Lemma 1 in B.1 for a formal proof).

We assume that all individuals have a strictly concave and twice differentiable utility function u, with $u' > 0$ and $u'' < 0$.

To characterize the set of Pareto efficient transfers under the local information constraint, we solve the following problem:

$$\max_{t \in \mathcal{T}^*} \mathbb{E}\left[\sum_{k \in N} \lambda_k u_k \left(e_k - \sum_{h \in N_k} t_{kh}\right)\right] \tag{5}$$

Note that each element in $\mathcal{T}^*$ corresponds to an "equivalent class" of transfer arrangements that are indistinguishable under the norm induced by $\langle \cdot, \cdot \rangle$. Moreover, as will be clear in the next three paragraphs, measure-0 changes in the transfers do not affect the objective functions (5). Hence, throughout the paper, we write "$s = t$" to mean "$\langle s - t, s - t \rangle = 0$", or equivalently "$s(\omega) = t(\omega)\ a.s.(\mathbb{P})$", and make statements with the understanding that the "$a.s.(\mathbb{P})$" qualifier applies whenever necessary.

Since the transfer rule t_{ij} is restricted to be measurable with respect to $\sigma(I_{ij})$, we may, with an abuse of notation, write it as $t_{ij} : \mathbb{R}^{\dim(I_{ij})} \to \mathbb{R}$. The following proposition provides a formal characterization of the solution to the maximization problem above.

Proposition 1. *A profile of $t \in \mathcal{T}^*$ solves (5) if and only if, for almost all possible states of world $\omega \in \Omega$, it simultaneously solves the $\sum_{i \in N} d_i$ optimization problems in the form of (6) at each common information set of the linked pair I_{ij}:*

$$t_{ij}(I_{ij}) \in \arg\max_{\tilde{t}_{ij} \in \mathbb{R}} \mathbb{E}\left[\lambda_i u_i \left(e_i - \tilde{t}_{ij} - \sum_{h \in N_i \setminus \{j\}} t_{ih}\right) + \lambda_j u_j \left(e_j + \tilde{t}_{ij} - \sum_{h \in N_j \setminus \{i\}} t_{jh}\right) \middle| I_{ij}\right] \tag{6}$$

$$\forall i, j \ s.t. \ G_{ij} = 1 \quad a.s.(\mathbb{P})$$

Proposition 1 is an intuitive result, and the analogue of it is easy to show when risk-sharing arrangements can be conditioned on everyone's income realization, as in Wilson (1968). In that case, for each possible realizations of $\omega \equiv (e_k)_{k \in N}$, we may freely choose $(t_{ij}(\omega))_{G_{ij}=1}$, a finite dimensional vector, to maximize (5). This enables for standard finite dimensional optimization techniques, leading to first order conditions for optimum that just connect ratios of marginal utilities of consumption for two individuals in two different states. In contrast, the intertwined local information structure induced by the locality constraint makes the optimization problem fundamentally infinite-dimensional. State-by-state optimization is no longer feasible: taking for example again the network depicted in Figure 1, we see that the local information set for the pair 12 is defined by $\{\tilde{\omega} \in \Omega : \ e_1(\tilde{\omega}) = e_1(\omega), \ e_2(\tilde{\omega}) = e_2(\omega)\}$ for a given ω, and the transfer t_{12} must be constant over all values of ω on I_{12}. As the value of t_{13} can vary continuously as a function of e_3 for each $\omega \in I_{12}$, the efficient t_{12} must be optimal in expectation with respect to the distribution of t_{13} conditional on I_{12}. Similarly, t_{13} must be optimal with respect to the conditional distribution of t_{12} on each I_{13}. As the realizations of I_{12} and the realizations of I_{13} induce two different uncountable partitions of the state space, with the transfers being constant on respective cells of the partitions, we can no longer carry out state-by-state optimization at each ω, but have to optimize the transfers at all states simultaneously.

The proof of Proposition 1 formally establishes this using mathematical results from infinite-dimensional convex optimization. We begin by establishing that under the inner product defined above, $\mathcal{T}^*$ forms a Hilbert space. We then show that the objective function is concave and twice Fréchet-differentiable. Then, the sufficient and necessary condition for optimality is given by the first Fréchet-derivative of the objective being the zero function, which is the infinite-dimensional generalization of the usual first-order condition for optimality.

(NONCOOPERATIVE FOUNDATION) Proposition 1 can also be regarded as a decentralization result for the social optimization problem. Notice that in problem (6), at each I_{ij}, the choice of t_{ij} affects the expected utilities of only i and j, so each optimization problem in (6) can be reinterpreted as the surplus maximization problem jointly solved by the linked pair ij, given the transfer rules chosen by other linked pairs. Then, the system of optimizations solved by all pairs in (6) can be regarded as a notion of equilibrium: each linked pair optimally chooses the risk-sharing arrangement along this link, given other risk-sharing arrangements along other links. Hence, Pareto efficiency on the social level is achieved whenever each pair makes efficient (defined for this pair only) choice on transfer rule.

The next result establishes that while in general there can be multiple transfer profiles satisfying the conditions for optimality (6), they all imply the same consumption plan in all states.

Proposition 2. *All profiles of transfers* $t \in \mathcal{T}^*$ *that solve (5) lead to (*$\mathbb{P}$*-almost) the same consumption plan* x*, where* $x_i(\omega) := e_i(\omega) - \sum_{j \in N_i} t_{ij}(\omega)$.

By Proposition 2, if we can find a profile of transfers so that the induced consumption plan satisfy (6), then it must correspond to a Pareto efficient risk-sharing arrangement.

For simplicity, below we will denote the conditional expectation operator $\mathbb{E}[\cdot \,|I_{ij}]$ by $E_{ij}[\cdot]$. In observation of Proposition 1 and Proposition 2, we may express the necessary and sufficient condition for Pareto efficiency as a requirement on the ratio of conditional expected marginal utilities given by the next Corollary.

Corollary 1. *A profile of transfers* t *is Pareto efficient if and only if the ratio of the expected marginal utilities conditional on all possible common information sets is constant: for every* $I_{ij} \in \mathcal{I}_{ij}$ *and every* $i, j \in N$ *s.t.* $G_{ij} = 1$,

$$\frac{E_{ij}\left[u_i'(x_i)\right]}{E_{ij}\left[u_j'(x_j)\right]} = \frac{\lambda_j}{\lambda_i}. \tag{7}$$

This extends the Borch rule (Borch (1962), Wilson (1968)) for Pareto efficient risk-sharing arrangements to settings with local information constraints. As opposed to the case when risk-sharing arrangements can be conditioned on the endowment realizations of all players, the ratio of expected marginal utilities of consumptions among individuals do not have to be equal state by state, they only have to be equal between linked individuals in expectation, conditional on common information.

4 Efficient Risk-sharing Arrangements under the CARA-Normal Setting

In this section we investigate Pareto efficient risk-sharing arrangements, subject to local information constraints, under the assumption of CARA utilities and jointly normally distributed endowments.

Assumption 1. *Throughout the subsequent sections we assume that individuals have homogeneous CARA utility functions* $u(x) = -e^{-rx}$*, where* $r > 0$ *is the coefficient* *of absolute risk aversion. The vector of endowments $(e_i)_{i\in N}$ follows a multivariate normal distribution, $e \sim N(\mathbf{0}, \sigma^2\Sigma)$ with, for some $\rho \in \left[-\frac{1}{n-1}, 1\right]$,*[10]

$$\Sigma := \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & 1 \end{pmatrix}.$$

4.1 Independent Endowments

We first analyze the case where endowments are independent, i.e., $e \sim N(\mathbf{0}, \sigma^2 \cdot \mathbf{I}_n)$.

We use a guess and verify approach, and postulate that a linear transfer scheme, that is a scheme for which the transfer between any two connected individuals is a linear function of endowment realizations in the pair's joint information set, can achieve any Pareto efficient risk-sharing arrangement. Below we verify that the candidate linear risk-sharing arrangement is globally optimal subject to local information constraints using the expectational Borch rule (7).

Given a linear transfer scheme, the final consumption, conditional on I_{ij}, also follow normal distribution, so $E_{ij}\left[u_i'(x_i)\right] = re^{-r\left(E_{ij}[x_i] - \frac{1}{2}rVar_{ij}[x_i]\right)}$. Define the conditional certainty equivalent $CE\left(x_i\right|I_{ij}) := E_{ij}[x_i] - \frac{1}{2}rVar_{ij}[x_i]$. Then (7) can then be rewritten as

$$CE\left(x_i^*\right|I_{ij}) - \frac{1}{r}\ln\lambda_i = CE\left(x_j^*\right|I_{ij}) - \frac{1}{r}\ln\lambda_j. \tag{8}$$

The profile of transfer schemes t achieves Pareto efficiency if and only if (8) holds for every pair of ij such that $G_{ij} = 1$, i.e., the difference in the conditional certainty equivalents is constant at each intersection of the information sets of a linked pair.

We say a profile of transfer rules is *strictly bilateral* if $t_{ij}(\omega) = t_{ij}(e_i, e_j)$.

Proposition 3. *Given any profile of positive welfare weights $(\lambda_i)_{i\in N}$, there always exists a strictly bilateral Pareto efficient profile of transfer rules in $\mathcal{T}^*$ in the form of $t_{ij}^*(e_i, e_j) := \frac{e_i}{d_i+1} - \frac{e_j}{d_j+1} + \mu_{ij}^*$, for some $\mu_{ij}^* \in \mathbb{R}$, for each linked pair ij.*

[10] $-\frac{1}{n-1}$ is the lower bound for a global pairwise correlation in a n-person economy; mathematically, it is the smallest ρ such that the variance-covariance matrix is positive semi-definite. For any $\rho \in \left[-\frac{1}{n-1}, 1\right]$, the variance-covariance matrix is positive semi-definite.

Recall that, by Proposition 2, the Pareto efficient consumption plan is always unique. But for general networks, there might be multiple risk-sharing arrangements that are Pareto efficient. In particular, superfluous transfers, which can be either state-dependent or state-independent, may be freely added to a cycle of individuals in the network without changing the final consumptions. Therefore, in general the transfer scheme achieving a Pareto efficient risk-sharing arrangement is not unique. In Appendix B.2 we show that for tree networks the linear transfer scheme featured in Proposition 3 is the unique transfer arrangement that achieves a given Pareto efficient risk-sharing arrangement.

The efficient transfer $t^*_{ij}(e_i, e_j)$ subject to the locality constraint is composed of two parts: the state-contingent "local equal sharing rule" and the state-independent "insurance premium". Furthermore, the transfers between two connected individuals ascribed by the linear transfer scheme in Proposition 3 only depend on endowment realizations of the two of them, not of their common neighbors. That is, only bilateral information is required for efficient risk sharing with local information. Also, ex ante two linked individuals ij only need knowledge of the local network structure (in particular d_i and d_j) to compute and contract on the socially optimal transfer rule t^*_{ij}.

4.2 Correlated Endowments

We now turn to the case of correlated endowments with $\rho \neq 0$. To maintain analytical tractability, in Assumption 1 we assume a symmetric correlation structure, where any two individuals' endowments have a constant pairwise correlation coefficient $\rho \in \left[-\frac{1}{n-1}, 1\right]$. Equivalently, we are assuming that each individual's endowment can be decomposed additively into two independent components: a common shock and a idiosyncratic shock, i.e., $e_i = \sqrt{\rho}\tilde{e}_0 + \sqrt{1-\rho}\tilde{e}_i$, with $(\tilde{e}_k)_{k=0}^n \sim_{iid} N(0, \sigma^2)$.

Below we show that any Pareto efficient risk-sharing arrangement can be achieved through a linear transfer scheme. We first consider the case of minimally-connected networks, for notational simplicity, and to develop intuition. Notice that, under minimal connectedness, $I_{ij} = (e_i, e_j)$, so transfer t_{ij} must be strictly bilateral. Then, the local FOC for optimality can be written as

$$t_{ij} = \frac{1}{2}e_i - \frac{1}{2}e_j - \frac{1}{2r}\ln \mathbb{E}\left[e^{r\sum_{k\in N_i\setminus\{j\}} t_{ik}(e_i,e_k)} \middle| e_i, e_j \right] + \frac{1}{2r}\ln \mathbb{E}\left[e^{r\sum_{k\in N_j\setminus\{i\}} t_{jk}(e_j,e_k)} \middle| e_i, e_j \right] + \frac{1}{2r}\ln\frac{\lambda_j}{\lambda_i} \tag{9}$$

Postulating a linear transfer scheme of the form, $t_{ij}(e_i, e_j) = \alpha_{ij}e_i - \alpha_{ji}e_j + \mu_{ij}$ $\forall G_{ij} = 1$, we can substitute the postulated linear forms of t_{ik} into (9) and obtain

$$\begin{aligned} t_{ij} = {} & \frac{1}{2}e_i - \frac{1}{2}e_j - \frac{1}{2}\sum_{k\in N_i\setminus\{j\}} \alpha_{ik}e_i - \frac{1}{2r}\ln \mathbb{E}\left[e^{-r\sum_{k\in N_i\setminus\{j\}} \alpha_{ki}e_k} \middle| e_i, e_j \right] \\ & + \frac{1}{2}\sum_{k\in N_j\setminus\{i\}} \alpha_{jk}e_j + \frac{1}{2r}\ln \mathbb{E}\left[e^{-r\sum_{k\in N_j\setminus\{i\}} \alpha_{kj}e_k} \middle| e_i, e_j \right] \\ & - \frac{1}{2}\sum_{k\in N_i\setminus\{j\}} \mu_{ik} + \frac{1}{2}\sum_{k\in N_j\setminus\{i\}} \mu_{jk} + \frac{1}{2r}\ln\frac{\lambda_j}{\lambda_i} \end{aligned}$$

Given the assumed correlation structure of endowments, the conditional distribution $e_k|_{e_i,e_j} \sim N\left(\frac{\rho}{1+\rho}(e_i+e_j),\ \frac{1+\rho-2\rho^2}{1+\rho}\cdot\sigma^2\right)$. We can explicitly derive the conditional expectation terms in the above formula and, after collecting terms and reconciling with the postulated formula for t_{ij}, we arrive at the following system of equations:

$$\alpha_{ij} = \frac{1}{2}\left[1 - \sum_{k\in N_i\setminus\{j\}} \alpha_{ik} + \frac{\rho}{1+\rho}\left(\sum_{k\in N_i\setminus\{j\}} \alpha_{ki} - \sum_{k\in N_j\setminus\{i\}} \alpha_{kj}\right)\right] \quad \forall ij \text{ s.t. } G_{ij} = 1 \tag{10}$$

In equation (10), the net transferred share α_{ij} of e_i from i to j is given by the half of the "remaining share" after deducting the transfers to i's other neighbors $N_i\setminus\{j\}$, corrected by an adjustment for inflows of non-local endowments. The $\frac{1}{2}$ multiplier is analogous to the equal sharing rule in the independent endowments case, but last term in the square brackets is new.[11] We refer to it as an informational effect, for the following reason. $\sum_{k\in N_i\setminus\{j\}} \alpha_{ki}$ is the sum of i's shares of i's other neighbors' endowments $(e_k)_{k\in N_i\setminus\{j\}}$, and the conditional expectation of each k's endowment changes linearly with the realization of e_i by a factor of $\frac{\rho}{1+\rho}$. Similarly, $\sum_{k\in N_j\setminus\{i\}} \alpha_{kj}$ is the sum of j's shares of j's other neighbors' endowments $(e_k)_{k\in N_j\setminus\{i\}}$, and the conditional expectation of each k's endowment also changes linearly with the realization of e_i

[11] This term disappears when $\rho = 0$.

by a factor of $\frac{\rho}{1+\rho}$. Due to the symmetric correlation structure, the realization of e_i provides the same amount of local information about all non-local endowment realizations e_k for $k \notin \overline{N}_i \cap \overline{N}_j$, and thus its informational effect can be calculated as a simple net summation of endowment shares. As a higher realized e_i predicts that both i and j are more likely to obtain higher amounts of inflows from uncommon neighbors, this commonly recognized information can be used by the pair ij to (imperfectly) share the non-local risk exposures.[12] After pooling the conditional expectations of non-local inflows, i and j again share the *remaining* shares of e_i and e_j equally. It is worth pointing out that i carries out this kind of "equal sharing" with all her neighbors, and the inflow/outflow shares $\{\alpha_{ij}\}$ must make all the sharing simultaneously equal (in expectation).

Hence, the $\left(\sum_{i\in N} d_i\right)$-dimensional vector $(\alpha_{ij})_{G_{ij}=1}$ must solve the system of $\left(\sum_{i\in N} d_i\right)$ linear equations defined by (10). No more linear restrictions need to be imposed on $(\alpha_{ij})_{G_{ij}=1}$, because, for each i, i herself absorbs $1-\sum_{j\in N_i}\alpha_{ij}$ of her own endowment so that e_i is fully shared within i's extended neighborhood. It can be shown that this system has a solution, but we leave this to be established later, for general networks.

For general network structure, the analysis is very similar to the above, but there are several complications. As $I_{ij} = \left(e_i, e_j, e_{N_{ij}}\right)$, the transfer rule t_{ij} can be contingent on $e_{N_{ij}} := (e_k)_{k\in N_{ij}}$ in addition to e_i, e_j. Furthermore, as the knowledge of the ex post realization of $e_{N_{ij}}$ brings in extra information about the distribution of non-local endowment realizations, Pareto efficiency requires that t_{ij} be contingent on $e_{N_{ij}}$. Specifically,

$$e_k|_{e_i,e_j,e_{N_{ij}}} \sim N\left(\frac{\rho}{1+(d_{ij}+1)\rho}\left(e_i+e_j+\sum_{k\in N_{ij}} e_k\right),\ V_{d_{ij}+2}\right) \tag{11}$$

where $d_{ij} := \#(N_{ij})$ and $V_{d_{ij}+2}$ denotes the variance of e_k conditional on observing $(d_{ij}+2)$ endowment realizations.[13]

We again postulate a linear transfer rule: $t_{ij} = \alpha_{ij}e_i - \alpha_{ji}e_j + \sum_{k\in N_{ij}}\beta_{ijk}e_k + \mu_{ij}$.

[12] To be precise, by "inflow" we mean the undertaking of a share of someone else's income endowment, which may be positive or negative; by "outflow" we mean the distribution of a share of one's own endowment to someone else, which may also be positive or negative. In particular, a negative inflow is not the same as an outflow. Instead, i's inflow from j is the same as j's outflow to i.

[13] See, for example, Eaton (2007), p116-117.

After some tedious algebraic transformations , we again arrive at a rather complicated system of linear equations in (α, β) that defines the condition for Pareto efficiency, namely system (25) (See Lemma 7 in Appendix A and B.1 for its explicit expression). However, instead of solving for this complicated system directly, we first present an innocuous simplification of it. Due to the possible existence of cycles and superfluous transfers along cycles, this system may in general admit multiple solutions. For example, given a complete triad ijk, we can make a superfluous transfer of a ϵ share of e_i from i to j, j to k and k to i by adding ϵ to α_{ij}, β_{jki}, and subtracting ϵ from α_{ik}. It can then be checked that this operation is indeed superfluous, in the sense that $(\alpha_{ij} + \epsilon, \beta_{jki} + \epsilon, \beta_{kji} - \epsilon, \alpha_{ik} - \epsilon)$, keeping everything else fixed, still solves the system of equations for Pareto efficiency with the induced final consumption plan left unchanged. Since any amount of superfluous cycles are redundant, we can set $\beta_{ijk} = 0$ for all triads ijk without loss of Pareto efficiency. Hence, in the following, we establish that there exists some vector of strictly bilateral transfer shares $(\alpha^*, \beta^* \equiv \mathbf{0})$ that solves (25) and thus achieves Pareto efficiency. In other words, the strictly bilateral linear transfer rules that we characterize below are the "simplest" Pareto efficient rules in terms of minimizing the sum of state-contingent transfers.

By setting $\beta = \mathbf{0}$, we achieve a significant simplification of (25) and obtain the following system:

$$\begin{cases} \alpha_{ij} = \frac{1}{2}\left(1 - \sum_{k \in N_i \setminus \{j\}} \alpha_{ik} + \gamma_{ij}\right) & (12.1) \\ 0 = \alpha_{ki} - \alpha_{kj} + \gamma_{ij} \quad \forall k \in N_{ij} & (12.2) \\ \gamma_{ij} = \frac{\rho}{1+(d_{ij}+1)\rho}\left(\sum_{k \in N_i \setminus \overline{N}_j} \alpha_{ki} - \sum_{k \in N_j \setminus \overline{N}_i} \alpha_{kj}\right) & (12.3) \end{cases} \forall i, j \text{ s.t. } G_{ij} = 1 \tag{12}$$

The first equation (12.1) states that the share of e_i transferred from i to j is half of the remaining share after i's transfers to i's other neighbors plus the informational adjustment term between ij. With $\gamma \equiv 0$, which is implied by $\rho = 0$, α will be simply reduced to the local equal sharing rule. The second equation (12.2) requires that the difference in the shares of e_k undertaken by i and j is equal to the informational effect between ij, so that it is indeed optimal for ij to set $\beta_{ijk} = 0$. This confirms again that strict bilaterality ($\beta = \mathbf{0}$) is not an assumption, as (12.2) also incorporates the efficiency requirements for $\beta = \mathbf{0}$. The third equation (12.3) defines the auxiliary variable γ_{ij}. We interpret γ_{ij} as the net informational effect because it is the rate at which locally observed endowment realizations affect the pair ij's joint expectation of non-local endowments. Notice that γ_{ij} is the same across $k \in \overline{N}_{ij}$ because each element of $(e_k)_{k\in\overline{N}_{ij}}$ provides exactly the same amount of information to the linked pair ij for their joint inference on non-local endowments. Given α, $|\gamma_{ij}|$ is decreasing in d_{ij}, indicating that the magnitude of the informational effect (for any single endowment realization) is decreasing in the amount of local information. Below we proceed to show the existence and provide a closed-form characterization of a solution to (12).

We first prove that (12.2) are implied by (12.1) and (12.3). By differencing (12.1) for ki and for kj we get: $\alpha_{ki}-\alpha_{kj} = \gamma_{ki}-\gamma_{kj}$. Hence, in the presence of (12.1) equation (12.2) is equivalent to, for all triads ijk, $\gamma_{ij} + \gamma_{jk} + \gamma_{ki} = 0$. This is reminiscent of the Kirchhoff Voltage Law for electric resistor networks, which states that the sum of voltage differences across any closed cycle must sum to zero. It turns out that the Kirchhoff Voltage Law indeeds holds in our setting for any cycle in a general network.

Proposition 4. ***"Kirchhoff Voltage Law"**: $\forall\rho \in \left(-\frac{1}{n-1}, 1\right)$, if (12.1) and (12.3) admit a solution (α, γ), this solution also satisfy (12.2); furthermore, given any cycle $i_1 i_2 ... i_m i_1$, γ satisfies the "Kirchhoff Voltage Law" $\gamma_{i_1 i_2} + \gamma_{i_2 i_3} + ... + \gamma_{i_m i_1} = 0$.*

Intuitively, Pareto optimality requires that individuals i and j share equally the net difference in the conditional expectations of *nonlocal* inflow exposures (captured by γ_{ij}) by creating an opposite net difference in their *local* inflow exposures, as specified in equation (12.2). This adjustment guarantees the expectational Borch rule in equation (7), and therefore Pareto efficiency. To see this, notice that conditional expectation and variance of consumption will differ only by a constant across different local endowment realizations (I_{ij}). Together, this implies that conditional CE's differ only by a constant, as required.

Given the redundancy of (12.2) in the presence of (12.1) and (12.3), as established in Proposition 4, we may now conclude that any solution to the system consisting of (12.1) and (12.3) defines a linear and Pareto efficient profile of transfer rules in $\mathcal{T}^*$. However, the matrix defined by (12) is difficult to work with, because of which we switch to an equivalent formulation of the system that is more tractable.[14] Specifically, we establish next that the Pareto efficiency is equivalent to minimization of the sum of consumption variances among linear risk-sharing arrangements in $\mathcal{T}^*$, and then

[14]Given Proposition 4, from now on, by (12) we mean the system of linear equations defined by (12.1) and (12.3).

show that the Pareto efficient risk-sharing arrangement is closely related to a network statistic that aggregates all *even-length paths* for every household, weighted in a particular way.

Let α be a linear profile of transfer rules in $\mathcal{T}^*$, and consider the following optimization problem that minimizes the sum of each individual's consumption variance under the risk-sharing arrangements defined by α:

$$\min_{\alpha} \sum_{i \in N} Var\left[\left(1 - \sum_{j \in N_i} \alpha_{ij}\right) e_i + \sum_{j \in N_i} \alpha_{ji} e_j\right]. \tag{13}$$

Writing $\alpha_{ii} := 1 - \sum_{j \in N_i} \alpha_{ij}$, the minimization problem (13) is equivalent to

$$\min_{(\alpha_{ij}),(\alpha_{ii})} \sum_i \left(\sum_{j \in \overline{N}_i} \alpha_{ji}^2 + 2\rho \sum_{j,k \in \overline{N}_i, j<k} \alpha_{ji}\alpha_{ki} \right) s.t. \sum_{j \in \overline{N}_i} \alpha_{ij} = 1 \quad \forall i \in N.$$

Let $\tilde{\Lambda}_i$ be the Lagrange multiplier associated with i's outflow constraint $\sum_{j \in \overline{N}_i} \alpha_{ij} = 1$ and denote $\Lambda_i := \frac{\tilde{\Lambda}_i}{2(1-\rho)}$. It is then straightforward to check that the set of admissible shares α is convex and the objective function is convex in α (as the underlying variance-covariance matrix is positive definite). Then, taking the FOC for the Lagrangian, we have

$$\begin{cases} \alpha_{ji} = \Lambda_j - \frac{\rho}{1-\rho}\left(\alpha_{ii} + \sum_{k \in N_i} \alpha_{ki}\right) \quad \forall j \in \overline{N}_i, \forall i \in N & (14.1) \\ \sum_{j \in \overline{N}_i} \alpha_{ij} = 1 \quad \forall i \in N & (14.3) \end{cases} \tag{14}$$

This is a system of $(\sum_i d_i + 2n)$ equations in $(\sum_i d_i + 2n)$ variables (α, Λ).

In order to obtain a closed-form characterization of Pareto efficient linear transfer arrangements, we first show that if (14) admits a unique solution, then this unique solution also solves (12). This confirms that the solution to (14) indeed defines a Pareto efficient profile of transfer rules in $\mathcal{T}^*$. We then show that (14) admits a closed-form solution that characterizes the linear profile of transfers. In Section 4.3 we relate both of these solutions by defining two related centrality measures that can be used interchangeably to describe efficient transfers in our model.

Proposition 5. $\forall \rho \in \left(-\frac{1}{n-1}, 1\right)$, *if system* (14) *admits a unique solution, then the* *solution also solves system (12); i.e., a profile of linear and strictly bilateral transfer rules is Pareto efficient in $\mathcal{T}^*$ if it uniquely minimizes the sum of consumption variances among all profiles of linear and strictly bilateral transfer rules in $\mathcal{T}^*$.*

We now show that, for any given network, system (14) indeed admits a unique solution that can be expressed in closed form. The solution depends on the pairwise correlation ρ and on the positions of individuals in the network, and can be represented as a linear function of accumulated paths along the network.

Proposition 6. [15] *For any $\rho \in \left(-\frac{1}{n-1}, 1\right)$ and any network structure G, or for $\rho = -\frac{1}{n-1}$ and any G such that $\max_{i\in N} d_i < n-1$, there exists a unique solution to system (14) given by the following: $\forall i \in N,\ \forall j \in \overline{N}_i$,*

$$\alpha_{ji} = \Lambda_j - \frac{\rho}{1+\rho d_i} \sum_{k \in \overline{N}_i} \Lambda_k \tag{15}$$

where Λ_i is given by:

- *(Recursive representation):*

$$\Lambda_i = \frac{1}{d_i+1} \left(1 + \sum_{j \in \overline{N}_i} \sum_{k \in \overline{N}_j} \frac{\rho}{1+\rho d_j} \Lambda_k \right) \tag{16}$$

- *(Closed-form representation): writing $\Lambda = (\Lambda_i)_{i=1}^n$*

$$\Lambda = \left(\overline{D} - \overline{G}\Psi\overline{G}\right)^{-1} \mathbf{1}$$

  where $\overline{D}$ is a diagonal matrix with its i-th diagonal entry being $d_i + 1$, Ψ is a diagonal matrix with its i-th diagonal entry being $\frac{\rho}{1+\rho d_i}$, and $\overline{G} := G + \mathbf{I}_n$.

- *(Explicit representation): $\forall \rho \in [0, 1)$,*

$$\Lambda_i = \frac{1}{d_i+1} + \sum_{q\in\mathbb{N}} \sum_{j\in N} \sum_{\pi_{ij}\in\Pi_{ij}^{2q}} W\left(\pi_{ij}\right) \tag{17}$$

[15] See Appendix B.3 for Pareto efficient risk-sharing arrangements in the boundary cases of $\rho \in \left\{-\frac{1}{n-1}, 1\right\}$.

where $W(\pi_{ij})$, the weight of each path $\pi_{ij} = (i_0, i_1, i_2, \ldots i_q)$ of length q from i to j (i.e. $i_0 = i$ and $i_q = j$), is given by,

$$W(\pi_{ij}) := \frac{1}{d_{i_0}+1} \cdot \frac{\rho}{1+\rho d_{i_1}} \cdot \frac{1}{d_{i_2}+1} \cdot \frac{\rho}{1+\rho d_{i_3}} \cdot \ldots \cdot \frac{1}{d_{i_q}+1} \tag{18}$$

Proposition 6 states that the form in which the network determines the Pareto efficient linear transfer arrangements has to do with interaction at distance two (i.e. neighbors of neighbors). Intuitively, the network interaction terms in (14.1) define substitutability across the shares going to j. This implies that individuals with a common neighbor (at distance two) interact directly as shown in (14.1) . Proposition 6 compounds these direct interactions as they extend throughout the entire network structure, showing that the relevant network statistic accumulates *weighted even paths* for every individual.

This concludes our complete characterization of the Pareto efficient risk-sharing arrangements subject to the local information constraint with globally correlated endowments.

4.3 Properties of Efficient Risk Sharing with Local Information

4.3.1 Network Centrality and Efficient Transfers

Proposition 6 and Proposition 4 suggest two potential measures of network centrality that, as functions of the network structure G and the correlation parameter ρ, directly enter into the determination of the Pareto efficient transfer shares. In the following section we explicitly define these two centrality measures, provide interpretations for them, and examine the relationship between them.

Let us first focus on the centrality measure constructed in Proposition 6. In observation of the "Kirchhoff Voltage Law", $\forall \rho \neq 0$, there exists a function $V : N \to \mathbb{R}$ s.t.

$$V(i) - V(j) = \gamma_{ij} \equiv \frac{\rho}{1 + (d_{ij}+1)\rho} \left(\sum_{k \in N_i \setminus \overline{N}_j} \alpha_{ki} - \sum_{k \in N_j \setminus \overline{N}_i} \alpha_{kj} \right).$$

In particular, V can be constructed from γ in the following way. Fix any individual, say, individual 1, and normalize $V(1) = 0$. For any other individual $j \neq 1$, as the network is connected, there exists a path $1i_1...i_mj$ that connects 1 to j. Then we simply define $V(j) = \gamma_{1i_1} + ... + \gamma_{i_mj}$. By Proposition 4, V is well-defined.

Using terminology from electrical resistor networks, $V(i)$ is analogous to an "energy potential" for node i, and γ_{ij} is the "potential difference" or "voltage" between i and j. Given γ, the "current flows" α (or more precisely, its deviation from local equal sharing) are driven by γ according to (12.1). However, γ is simultaneously determined by the currents α according to (12.3). The multiplier $\frac{\rho}{1+(d_{ij}+1)\rho}$ in the Γ matrix captures how much a pair ij "discounts" nonlocal inflows given the amount $(d_{ij} + 2)$ and the quality (ρ) of local information. Despite the possible differences in this discount multiplier across linked pairs, the optimum is associated with a globally consistent assessment of each individual's "net position in nonlocal risk exposures", which is summarized by $V(i)$. From now on, we refer to $V(i)$ (or simply V_i) as i's "*in-potential centrality*".

The other centrality measure we consider is the Lagrange multiplier Λ_i in Proposition 6, which is a network statistic that aggregates all even-length paths for every household, weighted in some particular way given by (18). In contrast with the in-potential centrality V_i, which is based on net difference in inflow shares, Λ_i relates to net difference in outflow shares. Hence, we thereafter refer to Λ_i as the "*out-potential centrality*".

Notice that recursive expressions in the flavor of (16) are often found in the definition of centrality measures in network analysis. For instance, in their well-known work on strategic complementarity in networks, Ballester, Calvó-Armengol, and Zenou (2006) show that equilibrium actions depend on a similar recursive measure known as Bonacich Centrality. More recently, Banerjee et al. (2014) have sought to identify individuals in the network that are best placed to diffuse information on micro-credit opportunities in India. They find that participation is higher if those first informed have higher eigenvector centrality. [16] All of these measures can be expressed generically as

$$B_i = c + \gamma \sum_k g_{ik} B_k \tag{19}$$

for some constant c and with $|\gamma| < 1$. This expression essentially says that $i's$ mea-

[16] Already at the beginning of the internet boom, a number of algorithms surfaced that allowed users to rank websites by their significance in the broader world wide web network. Procedures such as PageRank and HITS algorithm also refined measures recursively throughout the network.

sure depends linearly on the sum of measures that are connected to i. However, there are two crucial distinguishing differences with respect to the out-potential centrality measure defined in (17). First of all, notice that equation (16) does not sum over all centrality measures of i's neighbors, but instead sums over those of $i's$ neighbors' neighbors. In other words, the out-potential centrality is defined recursively at distance two, not one. This is not unique in in network analysis. It appears in some work on vertex similarity by Jeh and Widom (2002), and it has also appeared in newer page-ranking algorithms, such as the HITS algorithm.[17] Secondly, notice that, in contrast to equation (19), the out-potential centrality does not weight all measures at distance two with a common parameter γ. Instead, they weighted by the degree of the household that serves as a bridge between them. So for example, imagine two households k and i are both linked to a third household l, but are not linked to each other. Then, $k's$ measure will enter the definition of $i's$ measure, weighed by the degree of l. The type of weighting scheme in equation (18) can be thought of in terms of the accumulated local interactions. Recall that indirect interactions only represent the concatenation of various direct interactions linked together by the network constraints, and the weights $\frac{\rho}{1+\rho d_k}$ capture all the households in a given path engaged in direct interactions and the weighs $\frac{1}{d_k+1}$ capture the connecting household's constraint. This weighting scheme marks a crucial distinction vis--vis other measures, in that additional paths does not guarantee an increase in households' network measure.

We now relate the two centrality measures defined above to the Pareto efficient transfer shares in the next proposition, which also help explain the terms "*in/out-*potential centralities".

Proposition 7. *Let α be a strictly bilateral Pareto efficient transfer shares. Write $\Lambda_{ij} := \Lambda_i - \Lambda_j$. Then, for any linked pair ij and any $k \in \bar{N}_{ij}$,*

(i) $\gamma_{ij} = \alpha_{kj} - \alpha_{ki}$

(ii) $\Lambda_{ij} = \alpha_{ik} - \alpha_{jk}$

(iii) $\Lambda_{ij} - \gamma_{ij} = \alpha_{ii} - \alpha_{jj}$

[17] In the HTIS algorithm, a webpage is given both an authority and a hubness score, with the property that a website's authority is determined by the sum of the hubness scores of other websites it links to, while a website's hubness is determined by the sum of the authorities of websites it is linked by. This implies that each one of this measures is defined recursively at distance two.

By construction, γ_{ij} is based on net differences in inflow shares while Λ_{ij} is based on net differences in outflow shares. If i has a larger potential centrality than j ($\gamma_{ij} > 0$), i gets larger net nonlocal exposures (from his noncommon neighbors $N_i \backslash \overline{N}_j$) than j does, and at Pareto optimum this net difference is shared between i and j through an opposite difference in net exposures to common friends: $\alpha_{ki} - \alpha_{kj} = -\gamma_{ij} < 0$, i.e. i typically gets a smaller local exposure (from a common neighbors $\overline{N}_{ij}$) than j does. A corresponding relationship also holds for the out-potential centrality. If i has a larger out-potential centrality than j ($\Lambda_{ij} > 0$), i gives out a larger share of his own shock locally (to a common neighbor in $\overline{N}_{ij}$) than j does: $\alpha_{ik} - \alpha_{jk} = \Lambda_{ij} > 0$ for $k \in \overline{N}_{ij}$, and therefore gives out a smaller share of his own shock nonlocally than j does. Also, the discrepancy between the potential centrality difference γ_{ij} and the out-potential centrality difference Λ_{ij} equals the difference in self-shock exposures $\alpha_{ii} - \alpha_{jj}$.

4.3.2 Network Centrality and Consumption Variance

We seek to establish a relationship between network centrality of an individual and her consumption variance, given a Pareto efficient risk-sharing arrangement. The fact that, in a Pareto optimal arrangement, neighbors equally share expected state-dependent realizations conditional on their common information does not imply that their ex ante consumption variance is equal. The transfer scheme that achieves the equalization of conditional expectations of the state-dependent part of the consumption plan has the feature that the neighbor with a higher exposure to non-common endowment shocks ends up with higher consumption variance. Mathematically this is because conditional expectations of uncommon shocks are equal to a constant $\frac{\rho}{1+(d_{ij}+1)\rho}$ times the sum of common shocks, and this constant is strictly smaller than 1. Hence, a unit increase in exposure to non-common shocks is compensated by less than a unit decrease in exposure to common shocks. This implies that individuals who are more central in risk sharing end up with a higher consumption variance in our model.

We illustrate the above relationships between network centrality, transfer flows, and consumption variances in the context of star networks. Let c denote the center individual, who is connected to $n-1$ peripheral individuals, and none of the peripheral individuals are connected to each other. We use p to refer to a generic peripheral individual.

It is straightforward to show that a linear risk-sharing arrangement achieving

Pareto efficiency subject to local information constraints specifies the following endowment shares to be transferred:

$$\alpha_{cp} = \frac{2+2\left(n-1\right)\rho}{n\left(2+n\rho\right)}, \quad \alpha_{pc} = \frac{1+\rho}{2+n\rho}, \quad \gamma_{cp} = \frac{\left(n-2\right)\rho}{2+n\rho}.$$

Given this risk-sharing arrangement, if the potential centrality of peripheral individuals is normalized to 0, the potential centrality of the center is $V\left(c\right) = \frac{n-2}{2+n\rho}$. The consumption plan is given by:

$$x_c = \frac{2\left(1+\rho\right)-\left(n-2\right)^2\rho}{n\left(2+n\rho\right)} \cdot e_c + \frac{1+\rho}{2+n\rho} \cdot \sum_{k\neq c} e_k + C_i$$
$$x_p = \frac{2+2\left(n-1\right)\rho}{n\left(2+n\rho\right)} e_c + \frac{1+\left(n-1\right)\rho}{2+n\rho} e_p + C_p.$$

Hence the difference in consumption variances satisfies

$$Var\left(x_c\right) - Var\left(x_p\right) = \frac{\left(n-2\right)\left(1+\left(n-1\right)\rho\right)\left(1-\rho^2\right)}{\left(2+n\rho\right)^2} \geq 0$$

with equality only at $\rho \in \left\{-\frac{1}{n-1}, 1\right\}$. In particular, $Var\left(x_c\right) - Var\left(x_p\right) \to \frac{1-\rho^2}{\rho}$ as $n \to \infty$, and hence the consumption variance of the center can be much higher than the consumption variance of a periphery individual when ρ is low and n is high.

Centrality in risk sharing, as implied via the Pareto efficient risk-sharing arrangements subject to local information constraints, is not equivalent to standard notions of centrality, such as degree or eigenvector centrality. However, on typical networks they are highly positively correlated, implying that our model predicts a positive relationship between these centrality measures and consumption variance. This contrasts with the predictions of the model in AMS, in which enforcement constraints limit the efficiency of risk-sharing arrangements. We illustrate this point numerically, estimating the correlation predicted by our model and by the model in AMS, between an individual's centrality and consumption variance, via simulated endowment realizations in two real-world village networks from India from two different databases, each randomly selected and provided us by the researchers who collected the data.[18]

[18]The first network was provided to us by Erica Field and Rohini Pande, who collected it from villages in the districts of Thanjavur, Thiruvarur and Pudukkotai (Tamil Nadu) in India. In a subset of the villages, complete within-village network data was collected by surveying all households. The second network is from data collected by Abhijit Banerjee, Arun Chandrasekhar, Esther Duflo and

Table 1: Correlation between Centralities and Consumption Variances

		(A) Field & Pande		(B) BCDJ	
	Capacity	Degree	Eigenvector	Degree	Eigenvector
AMS	0.5	-0.8943^{***}	-0.4585^{***}	-0.3552^{***}	-0.3155^{***}
	1.0	-0.6885^{***}	-0.2313^{***}	-0.2913^{***}	-0.2573^{***}
	1.5	-0.5430^{***}	-0.0898^{***}	-0.2843^{***}	-0.2511^{***}
Our Model		0.1994^{***}	0.1659^{***}	0.2084^{***}	0.0858

*** denotes statistical significance at 1%-level.

In both simulations, we randomly drew the endowment $e_i^{(t)}$ of each household according to the standard normal distribution for $T = 5000$ times: $\left\{e_i^{(t)}\right\}_{i,t} \sim_{iid} N(0,1)$. We assumed that all households have CARA utility functions with $\lambda = 1$. We then computed the final consumptions of each household under the equally-weighted Utilitarian optimal risk-sharing arrangement subject to local information constraints, using the results from subsection 4.1, and the sample variance of final consumptions for each household (note that the variance does not depend on the planner's weights). Following this we computed the sample correlation between degree/eigenvector centrality and consumption variance. Similarly, we computed the constrained efficient consumptions implied by the model in AMS, and the sample correlation between the centrality measures and consumption variance under three levels of capacity constraints (the maximum amount that can be transferred through any link, at any state): 0.5, 1 and 1.5. The results are summarized in Table 1. All results are highly statistically significant with p-values very close to zero, except for the correlation between consumption variance and eigenvalue centrality for one of the two datasets).[19]

Under the AMS model, we observe a negative correlation between centrality and consumption variance. In AMS, transfers along links are subject to capacity constraints. As a result, centrally located households tend to have a lower consumption variance, because capacity constraints are less likely to be binding for them locally, and for typical endowment realizations they end up pooling risk with a larger set of

Matthew Jackson in Karnataka, India (they collected complete within-village network data in 75 villages), used for example in the Banerjee, Chandrasekhar, Duflo, and Jackson (2014). From both datasets we received the network of financial connection for one randomly selected village with complete network data. From the original network we created the undirected "AND" network, that is, we defined a link between two households whenever both households indicated each other as a borrowing relationship. We excluded households that became isolated in the "AND" network.

[19] The p-values, calculated from standard t-tests against the null hypotheses of zero correlations, are at orders of magnitudes below 10^{-10}, except the case noted.

other households.[20] This holds for all capacity values we used in the simulations, but the relationship is more highlighted for relatively stricter capacity constraints.[21]

Under the model of the current paper, we observe the opposite sign: sample correlation between both degree and eigenvector centrality on the one hand, and consumption variance on the other hand is positive (as noted above, not significantly for eigenvalue centrality when using the BCDJ data).

5 Implications for Empirical Tests of Risk Sharing

The performance of risk-sharing communities has been repeatedly tested in data since the work of Cochrane (1991), Mace (1991) and Townsend (1994). Their original approach developed empirical tests of full insurance that related household consumption and income. Indeed, the well known-Borch rule – equating the ratio of marginal utilities across households – imposes that, under full insurance, household consumption should not respond to idiosyncratic movements in income after controlling for aggregate shocks. This implication can be tested in the following popular regression:

$$log(c_{it}) = \alpha_i + \beta_1 log(y_{it}) + \beta_2 log(\bar{y}_t) + \epsilon_{it} \tag{20}$$

where c_{it} and y_{it} correspond to household $i's$ consumption and income at time t, and where $\bar{y}_t = \sum_i y_{it}$ represents aggregate village income at time t.[22] Full insurance implies that $\beta_1 = 0$ and $\beta_2 = 1$. An overwhelming proportion of studies have rejected the full-insurance hypothesis in a wide number of settings. As a result, a great deal of work has followed, that seeks to explain this stylized fact.

On the theory side, we have argued that this paper complements an ongoing effort to model the relevant contracting frictions in informal risk sharing environments.[23] In this section we argue that our framework also responds to a recent strand of the literature that suggests modifying the classical Townsend test in order to accommodate various forms of heterogeneity. Some of this work argues that the standard consump-

[20] Using terminology from AMS, more centrally-located households typically end up on larger "risk-sharing islands."

[21] As capacities increase, centrality in the AMS model matters less, since capacity constraints are less likely to bind.

[22] Village-time fixed effects are traditionally used to capture aggregate shocks at the village level.

[23] For example Thomas and Worrall (1990), Kocherlakota (1996), Ambrus et al. (2014), and Kinnan (2011).

tion regression in (20) is misspecified if, for instance, households hold heterogeneous risk preferences.[24] More relevant to the current discussion, several other studies have also suggested that households within a village indeed access different risk sharing groups, and that controlling for aggregate-level shocks, as in (20), would incorrectly estimate income coefficients: $\bar{y}$ should be group-specific. In a couple well-known examples, Mazzocco and Saini (2012) argue that the relevant sharing group in India is the caste (rather than the village), while Attanasio et al. (2015) test for efficient insurance within extended families in the U.S.[25]

This paper refines and generalizes the modified tests that evaluate the performance of insurance mechanisms on local sharing groups. Rather than taking groups as separate, perfectly insured communities, the current framework allows for a fully general social structure with interconnected sharing groups that are specific to each household, and which may overlap in complicated ways along any given network. We show how, under the local information constraints of our model, not defining the relevant local sharing group biases the estimates of risk-sharing tests. More importantly, we show that controlling for this bias will not eliminate the correlation between household consumption and income: the structure of the network, coupled with the information constraints, induces imperfect risk-sharing and generates heterogeneity in sharing behavior. The current framework therefore allows us to decompose the standard Townsend coefficient β_1 into an underlying distribution of household-specific coefficients that capture the varying risk-sharing possibilities induced by the network structure, and which can be interpreted economically in terms of consumption volatility (as shown in the previous section).

To fix ideas, consider the simple network with three individuals and independent endowments in section 2 and set $\lambda_i = 1$; all arguments below can be extended to general networks, correlated endowments, and any profile of Pareto weights. If we write down final consumption for each household in the form of the classical risk-sharing specification of equation (20), we have that,

[24] See for instance Mazzocco and Saini (2012) and Schulhofer-Wohl (2011).

[25] In similar procedures Hayashi et al. (1996) consider whether extended families can be viewed as collective units sharing risk efficiently. Munshi and Rosenzweig (2016) also find that the caste is the relevant group to explain migration patterns in rural India. Most relevant here, Fafchamps and Lund (2003) address the failure of efficient insurance in the data suggesting that households receive transfers not at the village level, but from a network of family and friends

$$\begin{cases} c_{1t} = \alpha_1 + \left(\frac{1}{3} - \frac{1}{2}\right) y_{1t} + \frac{1}{2}\bar{y}_t + \epsilon_{1t}, \\ c_{2t} = \alpha_2 + \left(\frac{1}{2} - \frac{1}{3}\right) y_{2t} + \frac{1}{3}\bar{y}_t + \left(\epsilon_{2t} - \frac{1}{3}y_{3t}\right), \\ c_{3t} = \alpha_3 + \left(\frac{1}{2} - \frac{1}{3}\right) y_{3t} + \frac{1}{3}\bar{y}_t + \left(\epsilon_{3t} - \frac{1}{3}y_{2t}\right), \end{cases}$$

where $\alpha_1 = \frac{1}{12}r\sigma^2$ and $\alpha_2 = \alpha_3 = \frac{1}{24}r\sigma^2$ correspond to state-independent transfers and are represented as household-specific intercepts. These equations reflect three important themes of this paper as they relate to empirical tests of risk-sharing: 1) coefficients on own income are generically different from zero for all households. i.e. $\alpha_{ii} \neq \alpha_{ij}$, 2) these coefficients vary according to households' network position, and 3) imposing the common sharing group on all households generates biased estimates: notice the last two equations contain weighted incomes in the error term. The classical risk sharing test in (20) pools these equations and obtains a unique estimate for β_1; given the previous discussion we expect this estimate to be biased, different from zero, and positive.

In order to obtain unbiased estimates for β_1, consider estimating (20) with the relevant local sharing group instead. In this case, we show coefficients are properly estimated, but we still obtain heterogeneous estimates, β_i, for the coefficients on own income. As a result, the risk sharing test still delivers positive estimates – not surprisingly, since risk sharing is not efficient under information constraints. To see this, rewrite again our consumption equations in the form of (20), but now allow for household-specific aggregates, $\bar{y}_{it} = \sum_{j\in N_i} y_{jt}$, that sum over the incomes of $i's$ sharing partners. In this case we have,

$$\begin{cases} c_{1t} = \alpha_1 + \left(\frac{1}{3} - \frac{1}{2}\right) y_{1t} + \frac{1}{2}\bar{y}_{1t} + \epsilon_{1t}, \\ c_{2t} = \alpha_2 + \left(\frac{1}{2} - \frac{1}{3}\right) y_{2t} + \frac{1}{3}\bar{y}_{2t} + \epsilon_{2t}, \\ c_{3t} = \alpha_3 + \left(\frac{1}{2} - \frac{1}{3}\right) y_{3t} + \frac{1}{3}\bar{y}_{3t} + \epsilon_{3t}, \end{cases}$$

Because aggregate income terms are now household-specific (i.e. $\bar{y}_i$), the additional terms in the error disappear and we obtain unbiased estimators. Notice, however, that coefficients to own income are different from zero so long as $\alpha_{ii} \neq \alpha_{ij}$. This implies that the pooled regression will again deliver positive coefficient for β_1, even with the appropriate local aggregates. In this context, the pooled estimate in fact represents the average of the underlying heterogeneity in risk-sharing possibilities across households, which respond to network effects and relate to consumption

Table 2: Simulated Risk-Sharing Test under the Model for Two Simple Economies

	Dependent variable: Consumption			
	Star Network		*Circle Network*	
	Common Group	Local Group	Common Group	Local Group
	(1)	(2)	(3)	(4)
Income	0.201	0.027	0.121	0.001
	(0.002)	(0.002)	(0.001)	(0.001)
Agg. Income	0.780	0.977	0.845	0.998
	(0.001)	(0.002)	(0.002)	(0.001)
Observations	300,000	300,000	400,000	400,000
R^2	0.529	0.692	0.477	0.654

Note: Income data simulated from log-normal distribution with $\sigma^2 = 4$ and $t = 100,000$
Model estimated on logged data with household-specific intercepts
Values in parentheses are standard errors

volatility as specified by the theoretical results above.

Finally, notice that under sufficiently symmetric structures, we cannot reject this localized version of the Townsend test, because in "regular" networks $\alpha_{ii} - \alpha_{ij} = 0$. This means we are able to generalize the discussion on appropriate local aggregates in Townsend regressions – the theory is sufficiently rich to accommodate previous models of local sharing groups, as well as many other local structures. In fact, a well-defined local version of the Townsend test may fail to reject full insurance not only if castes or extended families are perfectly connected partitions (as stressed in the previous literature), but also if the social structure is sufficiently symmetric. As an extreme example, consider the circle network in which all individuals are identically positioned. Although all local sharing groups overlap and none of them are perfectly connected, this network structure would nonetheless generate sufficient regularity to "pass" an appropriately defined version of the risk-sharing test.

The previous discussion can be observed compactly in table 2, where the risk-sharing test is performed on simulated income data for the three individual "star" network discussed above, and the four individual "circle" network that exhibits perfect symmetry. The test is performed both with a common aggregate income term (columns 1 and 3) and with appropriately defined local sharing groups (columns 2 and 4). Notice that coefficients on own income are biased upwards by a whole order of magnitude when imposing a common aggregate income term but remain positive and significant in the star network, where the lack of symmetry keeps the pooled coefficient estimate away from zero. However, as discussed above, the circle network "passes" the Townsend test (coefficient to income is not significant) under appropriately specified local aggregate income terms.

6 Discussion

6.1 Spatial Correlation Structure

In the previous section we considered a symmetric correlation structure, in which the correlation between the endowments of two individuals did not depend on their positions on the network. An alternative specification, however, is to incorporate the possibility of spatially correlated endowments, that is correlation that decays with social distance.[26] As we illustrate below (and in more details in Appendix B.4), this type of correlation structure can be detrimental to the efficiency of informal risk sharing with local information constraints.

For concreteness we assume that the correlation between e_i and e_j geometrically decays with the social distance between i and j : $Corr\left(e_i, e_j\right) = \varrho^{dist(i,j)}$, where the social distance $dist\left(i, j\right)$ is formally defined as the length (i.e., the number of links) of the shortest path connecting i and j in network G. Also, for analytical simplicity we focus on circle networks with $n = 2m + 1$ individuals. In order to make comparable the risk-sharing efficiencies under geometrically decaying spatial correlation structure with that under the uniform global correlation structure analyzed in Section 4, we control the "shareable risk" to be the same across the two specifications by setting $\rho = \rho_m\left(\varrho\right) := \frac{\varrho(1-\varrho^m)}{m(1-\varrho)}$, where ρ is the uniform global pairwise correlation, while ϱ is the rate of decay in the geometrically decaying correlation structure. Then informal risk sharing subject to the local information constraint achieves drastically different levels of asymptotic efficiency under the two correlation structures.

Proposition 8. *Let $x_i^{unif}\left(\rho\right), x_i^{geo}\left(\varrho\right)$ denote the Pareto efficient consumption plan subject to the local information constraint under the uniform and the geometrically decaying correlation structures, parametrized by ρ and ϱ respectively, and let $Var_{unif,\rho}$,*

[26] There are many reasons why this correlation structure is more realistic for certain types of endowment shocks: for example, as shown in Fafchamps and Gubert (2007b) and in Conley and Udry (2010), social distance tends to be highly correlated with geographic proximity.

$Var_{geo,\rho}$ correspond to the variance operators under the two probability distributions induced by the two correlation structures. Then:

$$\lim_{\varrho\to 1}\lim_{m\to\infty} Var_{unif,\rho_m(\varrho)}\left(x_i^{unif}\left(\rho_m\left(\varrho\right)\right)\right)=\frac{1}{3},$$
$$\lim_{\varrho\to 1}\lim_{m\to\infty} Var_{geo,\rho}\left(x_i^{geo}\left(\varrho\right)\right)=1.$$

Hence, for ϱ close to 1 and sufficiently large m, uniform correlation leads to significant risk sharing (yielding payoffs close to that under independent endowments), while geometrically decaying correlation yields payoffs very close to the autarky payoffs, even though the two correlation structures lead to the same payoffs if global information can be used for risk sharing.

This difference in risk-sharing efficiency, driven by the difference in underlying correlation structures, is a peculiar feature of the local information constraint considered in this paper. With global information, a geometrically decaying correlation structure does not in itself imply risk-sharing inefficiency relative to the uniform correlation structure. For example, in a large ring network considered above, shocks that are spatially far away from each other are almost independent, and each given individual is spatially far away from most of the individuals in the network. Hence, under global information mostly shocks with low correlations are pooled together, thus yielding significant risk reduction. However, with local information, only spatially close shocks are pooled, rendering risk sharing virtually ineffective due to the high local correlation.

This might help explain why it is the case that while in most settings empirical research found that informal insurance works well, Kazianga and Udry (2006) found a setting in which informal insurance does not seem to help, and Goldstein, de Janvry, and Sadoulet (2001) found that certain types of endowment shocks are not well insured through informal risk sharing. In particular, this may be due to high correlation between endowments of neighboring households in the above settings, for the types of endowment shocks investigated.

6.2 Alternative Model Specifications

The main results in Section 4 are developed under the CARA-normal setting (Assumption 1) with a global correlation structure. We now consider the extendability of those results under some alternative model specifications.

Quadratic Utility Function

As to the specification of utility functions, we could alternatively work with quadratic utility functions, $u_i(x_i) = x_i - \frac{1}{2}rx_i^2$ for $i \in N$, which also admits a mean-variance expected utility representation. Noting that $u_i^{'}(x_i) = 1 - rx_i$, the conditional Borch rule in Proposition 1 takes the form of $\lambda_i(1 - rE_{ij}[x_i]) = \lambda_j(1 - rE_{ij}[x_j])$. With equal Pareto weightings ($\lambda = \mathbf{1}$) and normal endowments, it can be shown that this leads to exactly the same system of linear equations as in (12).[27] Hence, the linear transfer shares given in Proposition 6 also characterize a Pareto efficient risk-sharing arrangement under the quadratic-normal setting. However, the Pareto efficient frontier traced out by all admissible Pareto weightings will correspond to a collection of different state-dependent transfer shares α.

Normality of Endowments Distribution

The family of normal distributions have two properties that are technically essential to the proof of Pareto efficiency via the conditional Borch rule. First, a linear combination of a jointly normal vector remains normal, which allows us to explicitly characterize the distribution of final consumption $x_i = e_i - \sum_{j \in N_i} \alpha_{ji} e_j$ when transfer rules are linear. Second, normal distributions admit linear conditional expectations in the form of (11), which allows us to transform the conditions for Pareto efficiency into a system of linear equations on transfer shares. The assumption of normality can be relaxed slightly: if the endowment vector has a joint elliptical distribution[28], then both properties carry over[29], and thus the transfer shares given by Proposition 6 continue to characterize the Pareto efficient risk-sharing arrangements. Without the joint normality (or ellipticity) assumption, linear risk-sharing arrangements are in generally not Pareto efficient. For example consider again the 3-individual line network with the random endowment vector $e = (Y, Z, -Z^3)$ where Y, Z are independent standard normal random variables. As there are effectively no uncertainty, the unique Pareto efficient profile of transfer rules is given by $t_{12}(e_1, e_2) = \frac{1}{3}e_1 - \frac{2}{3}e_2 - \frac{1}{3}e_2^3$, $t_{13}(e_1, e_3) = \frac{1}{3}e_1 - \frac{2}{3}e_3 - \frac{1}{3}e_3^{1/3}$ which are clearly nonlinear.

[27] The proof is available in Appendix B.5.

[28] Normal distribution is a special case of elliptical distribution.

[29] See, for example, Fang, Kotz, and Ng (1990), Theorem 2.16 & 2.18.

Heterogeneity in Expected Endowments

Throughout Section 4 we maintained the specification that endowment distributions have zero mean. However, we argue that, as risk sharing is the sole concern of this paper, the specification of zero mean is a warranted normalization. For concreteness, let $\overline{y}_i$ be the expected level of endowment for individual i, and $y_i = \overline{y}_i + e_i$ be the random realization of endowment, where e_i is assumed to have zero mean. Clearly y and e induce the same local information structures $\sigma\left(y_k : k \in \overline{N}_{ij}\right) \equiv \sigma\left(e_k : k \in \overline{N}_{ij}\right)$, so it makes no differences whether the risk-sharing arrangements are specified to be contingent on y or e. Hence our results remain valid regardless of whether "endowments" or "endowment shocks" are shared. Moreover, neither does it make any difference whether the linear "guess" is taken to be $t_{ij} = \alpha_{ij}e_i - \alpha_{ji}e_j + \sum_{k\in N_{ij}} \beta_{ijk}e_k + \mu_{ij}$ or $t_{ij} = \alpha_{ij}y_i - \alpha_{ji}y_j + \sum_{k\in N_{ij}} \beta_{ijk}y_k + \tilde{\mu}_{ij}$: both will lead to the same system of linear equations in (12), so the Pareto efficient state-dependent transfer shares are given by exactly the same formulas in Proposition 6, irrespective of the value of mean-income vector $\overline{y}$. Any difference induced by $\overline{y}$ is completely absorbed by the state-independent transfers μ, which are irrelevant to Pareto efficiency in our current framework.[30]

Local Observability of Endowments of kth-Order Neighbors

The specification of our model assumes that each individual can only observe endowments realizations of individuals at most one link away from him. Alternatively, we could assume local observability of endowment realizations at most k links away, and our model can be easily adapted to characterize the Pareto efficient risk-sharing arrangements under this relaxed observability assumption. Specifically, given a connected "physical network structure" G, we may define the "informational network structure" $\tilde{G}$ by adding a link between every pair of individuals with no larger a graph distance than k in G. The Pareto efficient consumption plan is then induced by the state-dependent transfer shares given by Proposition 6 with $\tilde{G}$ in place of G, as now $\tilde{G}$ induces the appropriate local information structure for the analysis in Section 3 and 4. The only concern lies in the characterization of a Pareto efficient profile of "physical" transfer rules that achieve the Pareto efficient consumption plan, because though we assume observability of kth-order neighbors, we may need to maintain

[30] In fact, the Pareto efficiency frontier traced out by all admissible Pareto weightings λ is independent of $\overline{y}$.

that transfers should only be carried out along physical links in G. As $\tilde{G}$ cannot be a minimally connected network, there is always multiplicity in Pareto efficient transfer profiles. In particular, it can be shown that there always exists a profile of transfer rules that achieves the Pareto efficient consumption plan given the informational network $\tilde{G}$ while respecting the physical network structure G. This can be constructed easily by letting direct neighbors acting as intermediaries to distribute the appropriate shares of farther away neighbors' shocks (within a graph radius of k).

Ex Post Communication

We now discuss briefly the potential impacts of various communication protocols on the set of contracts that are ex post implementable. Consider allowing for a single round of communication after endowment realizations, where each individual i can send a message m_{ij} to each other individual j. The risk-sharing contract between a linked pair may be specified as a "messaging mechanism" that, at each the realized local state, maps the locally commonly observable messages to an amount of net transfer. In Appendix B.6, we consider four simple yet natural communication protocols: (a) "global communication", (b) "local announcement", (c) "local comment" and (d) "private communication", each specifying a particular degree of observability of the messages sent by each individual. We then show that with (a) the global information first best is ex post Nash implementable, with (b)(c) some improvements in risk-sharing efficiency are implementable, but with (d) no improvement is implementable. In each case, the Pareto efficient consumption plan with communication is consistent with our result in Proposition 6 applied to an augmented "informational network", a supergraph of the underlying "physical network" G specific to the communication protocol. In this sense, our earlier result may be interpreted as a characterization of Pareto efficient risk sharing for any given well-defined "informational network". A more rigorous analysis of how informational networks arises from physical networks under different communication protocols is left for further investigation.

6.3 State-Independent Transfers

In the previous sections we primarily focused on state-dependent transfer shares α, the only relevant variable under the CARA-normal setting as far as Pareto efficiency is concerned. We now provide a brief analysis of the state-independent constant transfers μ. Fixing any admissible Pareto weightings λ and any vector of expected endowments $\overline{y}$, there will be a profile of state-independent transfers μ^* associated with a solution to the the social planner's problem. μ^* can be in general decomposed into three components: $\mu^*_{ij} = \hat{\mu}_{ij}(\sigma^2) + \overline{\mu}_{ij}(\overline{y}) + \tilde{\mu}_{ij}(\lambda)$, with $\hat{\mu}_{ij}(0) = 0$, $\overline{\mu}_{ij}(\overline{y}) = 0$ if $\overline{y} \in Span(\mathbf{1})$, and $\tilde{\mu}_{ij}(\lambda) = 0$ if $\lambda \in Span(\mathbf{1})$. The first term $\hat{\mu}_{ij}$ admits a clear economic interpretation as the "insurance premium": individuals at different network positions undertake different exposures of endowment shocks and linked pairs typically exchange shocks of different sizes, so those who are subject to relatively larger final consumption variance are compensated through $\hat{\mu}_{ij}$. This component of constant transfers is of economic significance in real risk-sharing environments, as it relates to the division of risk-sharing surpluses and thus affect incentives in endogenous network formation. (See the next subsection for a more detailed analysis). The second term $\overline{\mu}$, however, should be interpreted as "inequality reduction", as this term serves equalize consumption even if there is no risk to share at all (when σ^2 is set to be zero), and is thus completely irrelevant to the focus of this paper. Hence this term should be always set to zero if risk sharing is the sole concern, and in this sense risk sharing via informal transfers should be better interpreted as the sharing of "endowment shocks" e, which suffice for the derivation of $\hat{\mu}$. This also partially motivated our focus on the case of $\overline{y} = \mathbf{0}$ in addition to the remark made in the last subsection. Lastly, the third term $\tilde{\mu}$is driven by the arbitrariness in welfare weightings: throughout the paper this remains as a auxiliary variable mathematically and does not admit a clear economic interpretation either.

6.4 Endogenous Network Formation

So far our analysis focused on characterizing Pareto efficient risk-sharing arrangements subject to local information constraints on an exogenously given network, implicitly assuming that the network structure is mainly shaped by predetermined factors such as kinship. Here we briefly discuss some implications of allowing for endogenous link formation in the context of informal risk sharing with local information constraints. The approach we take is similar as in Ambrus, Chandrasekhar, and Elliott (2015), who consider network formation in a risk-sharing framework with global information contracts, and propose a two-stage game in which in the first stage individuals can simultaneously indicate other individuals they want to link with. If two individuals each indicated each other, the link is formed, and the two connecting individuals each incur a cost of $c \geq 0$.[31] The solution concept we use is pairwise stability. In the second stage, whatever network is formed in the first stage, it is assumed that individuals agree on a Pareto efficient risk-sharing arrangement subject to local information constraints.

In our analysis of the CARA-normal framework so far, state independent transfers played a very limited role. However, when we allow for endogenous network formation, it becomes crucial how the network structure influences state independent transfers, and hence the distribution of surplus created by risk sharing, as it directly affects incentives to form links. Therefore, it is important to specify exactly which Pareto efficient risk-sharing arrangement prevails for each possible network that can form. Different ways of specifying state-independent transfers can lead to very different conclusions regarding network formation, as we demonstrate below.

A benchmark case is when all state-independent transfers are set to 0, which case is extensively investigated by Gao and Moon (2016) who assume local equal sharing with no state-independent transfers as an ad hoc sharing rule. They show that, even with zero cost of linking, an individual i's benefit for establishing an extra link with j falls very fast with the existing number of links the individual i has, as with more existing neighbors (larger d_i) the marginal reduction in self-endowment exposure $\left(\frac{1}{d_i+1} - \frac{1}{d_i+2}\right)$ is small relative to the additional exposure to j's endowment $\frac{1}{d_j+2}$. Typically this implies severe underinvestment into social links.

An alternative approach is pursued by Ambrus, Chandrasekhar, and Elliott (2015), in the context of risk-sharing arrangements with global information: they assume that the profile of state-independent transfers is determined according to the Myerson value. The Myerson value, proposed in Myerson (1980), is a network-specific version of the Shapley value that allocates surplus according to average incremental contribution of individuals to total social surplus.[32] Ambrus, Chandrasekhar, and Elliott (2015) in particular show that with state-independent transfers specified as above (for whatever network is formed), if individuals are ex ante symmetric then there is never underinvestment, that is given any stable network, there is no potential

[31] This simple game of network formation was originally considered in Myerson (1991). See also Jackson and Wolinsky (1996).

[32] Ambrus, Chandrasekhar, and Elliott (2015) also provide micro-foundations, in the form of a decentralized bargaining procedure between neighboring individuals that leads to state independent transfers achieving the Myerson value allocation.

link that is not established, even though its net social value would be strictly positive. Below we show that the same conclusion holds in our setting with local information constraints, in the case of CARA utilities and independently an jointly normally distributed endowments. The detailed specification and the proof are available in Appendix B.7.

Proposition 9. *Suppose that, for any given network structure, the Pareto efficient consumption plan subject to the local information constraint is implemented, and the state-independent transfers are induced by the Myerson values. Consider the first-stage network formation game in which each individual pays a private cost of c for each of her established links. Then, there is no underinvestment in social links in any pairwise stable network.*

We leave a more detailed investigation of network formation in the context of risk sharing with local information constraints to future research.

7 Conclusion

This paper analyzes informal risk sharing arrangements assuming that only direct neighbors may observe each others' endowment realizations, and that bilateral transfers must depend on commonly observed information only. Relative to previous models that propose alternative explanations for the inefficiency of these informal insurance arrangements, our framework provides a number of new and testable predictions. First of all, the model provides closed-form expressions for the set of bilateral exchanges that obtain a Pareto efficient allocation, for any network structure. We find that centrally located individuals become quasi insurance providers to more peripheral households. Further, the current setup formalizes, and indeed generalizes, the notion of a "local sharing group" that has been invoked recently in the risk-sharing tests performed in the development literature. We also show that the performance of these risk sharing arrangements is highly sensitive to the type of correlation structures that we assume, and in particular to the assumption that correlations decrease over network distance. This can potentially help explain why informal risk sharing works better in certain settings than in others.

The model provides numerous implications for empirical work. In a first approach, Milán (2016) shows that the current framework fits the observed sharing behavior of indigenous communities in the Bolivian Amazon. However, further empirical work is needed to distinguish local information constraints from other similar contractual frictions, such as the hidden income model identified by Kinnan (2011) as the relevant friction in Thai data. Indeed, an exciting new project seeks to derive a dynamic version of the model that provides testable predictions between current consumption and past information, which can be compared to those of other proposed risk-sharing frictions. Another empirical project to follow from this work takes the model's predictions on bilateral exchanges in order to develop a complete model of spillover effects across individuals that can be used to structurally estimate the underlying network structure following techniques in Manresa (2016). Finally, another interesting extension of the model considers the possibility of endogenous network formation. Following the work of Ambrus, Chandrasekhar, and Elliott (2015), this extension introduces interesting strategic effects that anticipate the type of constrained efficient transfers accrued to each pair, as predicted in the current model.

References

Ambrus, A., A. Chandrasekhar, and M. Elliott (2015): "Social Investments, Informal Risk Sharing, and Inequality," *Working Paper*.

Ambrus, A., M. Mobius, and A. Szeidl (2014): "Consumption Risk-Sharing in Social Networks," *The American Economic Review*, 104, 149–182.

Angelucci, M., G. de Giorgi, and I. Rasul (2015): "Resource Pooling within Family Networks: Insurance and Investment," Working Paper 2649, eSocialSciences.

Attanasio, O., C. Meghir, and C. Mommaerts (2015): "Insurance in Extended Family Networks," Working Paper 21059, NBER.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's Who in Networks. Wanted: the Key Player," *Econometrica*, 74, 1403–1417.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2014): "Gossip: Identifying Central Individuals in a Social Network," *National Bureau of Economic Research Working Paper*.

Blanchard, P. and E. Brüning (2012): *Mathematical Methods in Physics: Distributions, Hilbert Space Operators, and Variational Methods*, Springer Science & Business Media.

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

Borch, K. (1962): "Application of Game Theory to Some Problems in Automobile Insurance," *Astin Bulletin*, 2, 208–221.

Bourlès, R., Y. Bramoullé, and E. Perez-Richet (2016): "Altruism in Networks," *Working Paper*.

Bramoullé, Y. and R. Kranton (2007): "Risk-Sharing Networks," *Journal of Economic Behavior & Organization*, 64, 275–294.

Coate, S. and M. Ravallion (1993): "Reciprocity without Commitment: Characterization and Performance of Informal Insurance Arrangements," *Journal of Development Economics*, 40, 1–24.

Cochrane, J. (1991): "A simple test of consumption insurance," *Journal of Political Economy*, 99, 957–976.

Conley, T. G. and C. R. Udry (2010): "Learning about a New Technology: Pineapple in Ghana," *The American Economic Review*, 35–69.

Deaton, A. (1992): "Household Saving in LDCs: Credit Markets, Insurance and Welfare," *The Scandinavian Journal of Economics*, 253–273.

Dubois, P., B. Jullien, and T. Magnac (2008): "Formal and informal risk sharing in LDCs: Theory and empirical evidence," *Econometrica*, 76, 679–725.

Eaton, M. L. (2007): *Multivariate statistics: a vector space approach*, Institute of Mathematical Statistics.

Fafchamps, M. (1999): "Risk sharing and quasi-credit," *Journal of International Trade & Economic Development*, 8, 257–278.

Fafchamps, M. and F. Gubert (2007a): "The formation of risk sharing networks," *Journal of development Economics*, 83, 326–350.

——— (2007b): "Risk sharing and network formation," *The American economic review*, 97, 75–79.

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

Fang, K.-T., S. Kotz, and K. W. Ng (1990): *Symmetric multivariate and related distributions*, Chapman and Hall.

Gao, W. Y. and E. Moon (2016): "Informal Insurance Networks," *The B.E. Journal of Theoretical Economics*, 16, 455484.

Goldstein, M., A. de Janvry, and E. Sadoulet (2001): "You cant always get what you want: Inclusion and Exclusion in Mutual Insurance Networks in Southern Ghana," .

Grimard, F. (1997): "Household consumption smoothing through ethnic ties: evidence from Cote d'Ivoire," *Journal of Development Economics*, 53, 391–422.

Hayashi, F., J. Altonji, and L. Kotlikoff (1996): "Risk-Sharing between and within Families," *Econometrica*, 64, 261–294.

Hooper, P. (2011): "The structure of energy production and redistribution among Tsimane' forager-horticulturalists," Phd dissertation, University of New Mexico.

Jackson, M. O. and A. Wolinsky (1996): "A strategic model of social and economic networks," *Journal of economic theory*, 71, 44–74.

Jeh, G. and J. Widom (2002): "SimRank: a measure of structural-context similarity," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 538–543.

Karaivanov, A. and R. M. Townsend (2014): "Dynamic financial constraints: Distinguishing mechanism design from exogenously incomplete regimes," *Econometrica*, 82, 887–959.

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

Kazianga, H. and C. Udry (2006): "Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso," *Journal of Development Economics*, 79, 413–446.

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

Kocherlakota, N. R. (1996): "Implications of efficient risk sharing without commitment," *The Review of Economic Studies*, 63, 595–609.

Ligon, E. (1998): "Risk sharing and information in village economies," *The Review of Economic Studies*, 65, 847–864.

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

Mace, B. (1991): "Full Insurance in the Presence of Aggregate Uncertainty," *Journal of Political Economy*, 99, 928–956.

Manresa, E. (2016): "Social Investments, Informal Risk Sharing, and Inequality," Working paper, MIT.

Mazzocco, M. and S. Saini (2012): "Testing Efficient Risk Sharing with Heterogeneous Risk Preferences," *American Economic Review*, 102, 428–68.

Milán, P. (2016): "Network-Constrained Risk Sharing in Village Economies," Working Paper 912, Barcelona GSE.

Munshi, K. and M. Rosenzweig (2016): "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, 106, 46–98.

Myerson, R. B. (1980): "Conference structures and fair allocation rules," *International Journal of Game Theory*, 9, 169–182.

——— (1991): *Game theory: analysis of conflict*, Cambridge,MA: Harvard University.

Paxson, C. H. (1993): "Consumption and income seasonality in Thailand," *Journal of political Economy*, 39–72.

Rosenzweig, M. R. (1988): "Risk, implicit contracts and the family in rural areas of low-income countries," *The Economic Journal*, 98, 1148–1170.

Schulhofer-Wohl, S. (2011): "Heterogeneity and Tests of Risk Sharing," *The Journal of Political Economy*, 119, 925–958.

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. M. (1982): "Optimal multiperiod contracts and the gain from enduring relationships under private information," *The Journal of Political Economy*, 1166–1186.

——— (1994): "Risk and insurance in village India," *Econometrica: Journal of the Econometric Society*, 539–591.

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

Wang, C. (1995): "Dynamic insurance with private information and balanced budgets," *The Review of Economic Studies*, 62, 577–595.

Wilson, R. (1968): "The theory of syndicates," *Econometrica: journal of the Econometric Society*, 119–132.

Appendix A: Lemmas and Proofs

The proofs for all the lemmas stated below are available in Appendix B.1.

Define $J(t) := \mathbb{E}\left[\sum_{k\in N}\lambda_k u_k\left(e_k - \sum_{h\in N_k} t_{kh}\right)\right]$, the objective function in equation (5).

Lemma 1. *$\mathcal{T}^*$ with $\langle\cdot,\cdot\rangle$ forms a Hilbert space.*

Lemma 2. *J is concave on $\mathcal{T}^*$.*

Lemma 3. *J is twice Fréchet-differentiable.*

Lemma 4. *For any $t\in\mathcal{T}^*$ that solves (6), we have $J^{'}(t) = \mathbf{0}$.*

Lemma 5. *The set of consumption plan induced by the profiles of transfer rules t in $\mathcal{T}^*$ is convex.*

Proof of Proposition 1

Proof. We first prove the "only if" part. Note that, given any $t\in\mathcal{T}^*$, $\forall i,j$,

$$\begin{aligned}\mathbb{E}\left[\sum_{k\in N}\lambda_k u_k\left(e_k-\sum_{h\in N_k}t_{kh}\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{k\in N}\lambda_k u_k\left(e_k-\sum_{h\in N_k}t_{kh}\right)\middle| I_{ij}\right]\right]\\ &\leq \mathbb{E}\left[\max_{t_{ij}\in\mathbb{R}}\mathbb{E}\left[\sum_{k\in N}\lambda_k u_k\left(e_k-\sum_{h\in N_k}t_{kh}\right)\middle| I_{ij}\right]\right]\end{aligned}$$

This is because, conditional on I_{ij}, t_{ij} must be constant across all possible states, and thus the maximization of the conditional expectation is to solve for the optimal real number t_{ij}. For t to be a solution for problem (5), suppose there exists linked ij such that t_{ij} does not solve the problem (6). Then, by the inequality above, there exists another t_{ij}, specified for each different realization of I_{ij} and hence each possible state of nature, that leads to higher value of $\mathbb{E}\left[\sum_{k\in N}\lambda_k u_k\left(e_k-\sum_{h\in N_k}t_{kh}\right)\right]$, contradicting the optimality of t for problem (5). Note that the "$\mathbb{P}$-almost-all" quantifier applies here.

For the "if" part, notice that by Lemma 4, t solves all (6) simultaneously implies that $J^{'}(t) = \mathbf{0}$. As $\mathcal{T}^*$ is a Hilbert space by Lemma 1 and $J:\mathcal{T}^*\to\mathbb{R}$ is concave by Lemma 2 and twice Fréchet-differentiable by Lemma 3, we can apply a mathematical result on convex optimization in Hilbert space asserting that if $J^{'}(t) = \mathbf{0}$, then $J(t)$ is a local maximum.[33] Moreover, this local maximum must also be a global maximum. Suppose not. Then there exists $s \in \mathcal{T}^*$ s.t. $J^{'}(s) = 0$ and $J(s) > J(t)$. Then by the concavity of the objective function,

$$J(\alpha s + (1-\alpha)t) \geq \alpha J(s) + (1-\alpha) J(t) > J(t)$$

for all $\alpha \in (0,1)$, contradicting with the fact that $J(t)$ is a local maximum. $\square$

Proof of Corollary 1

Proof. By the concavity (shown in Lemma 2) of the objective function in (6), the FOC is both sufficient and necessary for maximization. The FOC w.r.t t_{ij}, is

$$\mathbb{E}\left[\lambda_i u_i^{'}\left(e_i - \sum_{h\in N_i} t_{ih}\right) + \lambda_j u_j^{'}\left(e_j - \sum_{h\in N_j} t_{jh}\right)\cdot(-1)\middle| I_{ij}\right] = 0$$

Rearranging the above we have

$$\frac{E_{ij}\left[u_i^{'}(x_i)\right]}{E_{ij}\left[u_j^{'}(x_j)\right]} = \frac{\mathbb{E}\left[u_i^{'}\left(e_i - \sum_{h\in N_i} t_{ih}\right)\middle| I_{ij}\right]}{\mathbb{E}\left[u_j^{'}\left(e_j - \sum_{h\in N_j} t_{jh}\right)\middle| I_{ij}\right]} = \frac{\lambda_j}{\lambda_i}.$$

$\square$

Proof of Proposition 2

Proof. Following the proof of Lemma 2, we can easily show, by the strict concavity of $u_i(\cdot)$, that the objective function in (5) is strictly concave in the consumption plan x. Lemma 5 shows that the set of admissible consumption plan induced by the set of transfer rules in $\mathcal{T}^*$ is convex. Hence, there is at most of one consumption plan that solves (5) . $\square$

Lemma 6. *Given any real vector $c \in \mathbb{R}^n$ such that $\sum_{i\in N} c_i = 0$, there exists a real vector $\mu \in \mathbb{R}^{\sum_i d_i}$ such that $\mu_{ik} + \mu_{ki} = 0$ for every linked pair ik and*

$$\sum_{k\in N_i} \mu_{ik} = c_i.$$

[33] See, for example, Theorem 30.2.2.(b) in Blanchard and Brüning (2012) on pp.394-395.

The solution is unique if and only if the network is minimally connected.

Proof of Proposition 3

Proof. Let x_i^* be the consumption plan induced by the transfer t^* described above. Then

$$\begin{aligned}
CE\left(x_i^*|I_{ij}\right) &= E_{ij}\left[e_i - \sum_{k\in N_i} t_{ik}^*\right] - \frac{1}{2} r Var_{ij}\left[e_i - \sum_{k\in N_i} t_{ik}^*\right] \\
&= e_i - \frac{e_i}{d_i+1} + \frac{e_j}{d_j+1} - \mu_{ij}^* - \sum_{k\in N_{ij}}\left(\frac{e_i}{d_i+1} - \frac{e_k}{d_k+1} + \mu_{ik}^*\right) \\
&\quad - \sum_{k\in N_i\setminus \overline{N}_j}\left(\frac{e_i}{d_i+1} - \frac{E_{ij}\left[e_k\right]}{d_k+1} + \mu_{ik}^*\right) - \frac{1}{2} r Var\left[\sum_{k\in N_i\setminus\overline{N}_j}\frac{e_k}{d_k+1}\right] \\
&= \frac{e_i}{d_i+1} + \frac{e_j}{d_j+1} + \sum_{k\in N_{ij}}\frac{e_k}{d_k+1} - \sum_{k\in N_i}\mu_{ik}^* - \frac{1}{2} r\sigma^2 \cdot \sum_{k\in N_i\setminus\overline{N}_j}\frac{1}{(d_k+1)^2}.
\end{aligned}$$

The necessary and sufficient condition for t^* to be Pareto efficient is given by (8). Plugging the above into (8) and canceling out the terms dependent on local information $(e_k)_{k\in\overline{N}_{ij}}$, we arrive at the following condition for Pareto efficiency:

$$\sum_{k\in N_i}\mu_{ik}^* + \frac{1}{2} r\sigma^2 \cdot \sum_{k\in N_i\setminus\overline{N}_j}\frac{1}{(d_k+1)^2} + \frac{1}{r}\ln\lambda_i = \sum_{k\in N_j}\mu_{jk}^* + \frac{1}{2} r\sigma^2 \cdot \sum_{k\in N_j\setminus\overline{N}_i}\frac{1}{(d_k+1)^2} + \frac{1}{r}\ln\lambda_j \tag{21}$$

Any profile of state-independent transfers μ^* that solves the above system (21) makes t^* efficient under weightings λ.

Notice that, if $CE\left(x_i^*\middle| I_{ij}\right) - \frac{1}{r}\ln\lambda_i = CE\left(x_j^*\middle| I_{ij}\right) - \frac{1}{r}\ln\lambda_j$ holds for any ω,

$$\begin{aligned}
CE\left(x_i^*\right) - \frac{1}{r}\ln\lambda_i &= \mathbb{E}\left[CE\left(x_i^*\middle| I_{ij}\right) - \frac{1}{r}\ln\lambda_i\right] - \frac{1}{2} r Var\left[CE\left(x_i^*\middle| I_{ij}\right) - \frac{1}{r}\ln\lambda_i\right] \\
&= CE\left(x_j^*\right) - \frac{1}{r}\ln\lambda_j
\end{aligned}$$

Hence, with G assumed WLOG to be connected, we have

$$CE\left(x_i^*\right) - \frac{1}{r}\ln\lambda_i = \frac{1}{n}\sum_{k\in N}\left(CE\left(x_k^*\right) - \frac{1}{r}\ln\lambda_k\right)$$

$$= -\frac{r\sigma^2}{2n}\sum_{k\in N}\frac{1}{d_k+1} - \frac{1}{nr}\sum_{k\in N}\ln\lambda_k \tag{22}$$

On the other hand, as $x_i^* = \frac{e_i}{d_i+1} + \sum_{k\in N_i}\left(\frac{e_k}{d_k+1} - \mu_{ik}^*\right)$,

$$CE\left(x_i^*\right) = -\sum_{k\in N_i}\mu_{ik}^* - \frac{1}{2}r\sigma^2\sum_{k\in\overline{N}_i}\frac{1}{\left(d_k+1\right)^2} \tag{23}$$

Equating the expressions for $CE\left(x_i^*\right)$ in (22) and (23), we obtain

$$\sum_{k\in N_i}\mu_{ik}^* = \frac{1}{2}r\sigma^2\left(\frac{1}{n}\sum_{k\in N}\frac{1}{d_k+1} - \sum_{k\in\overline{N}_i}\frac{1}{\left(d_k+1\right)^2}\right) + \frac{1}{r}\left(\frac{1}{n}\sum_{k\in N}\ln\lambda_k - \ln\lambda_i\right). \tag{24}$$

Lemma 6 has established that there indeed exists a solution μ^* to (24). Given any solution μ^* to (24), as $\overline{N}_i\backslash\left(N_i\backslash\overline{N}_j\right) = \overline{N}_{ij}$, we have

$$\begin{aligned}\sum_{k\in N_i}\mu_{ik}^* + \frac{1}{2}r\sigma^2\sum_{k\in N_i\backslash\overline{N}_j}\frac{1}{\left(d_k+1\right)^2} + \frac{1}{r}\ln\lambda_i &= \frac{1}{2}r\sigma^2\left(\frac{1}{n}\sum_{k\in N}\frac{1}{d_k+1} - \sum_{k\in\overline{N}_{ij}}\frac{1}{\left(d_k+1\right)^2}\right) + \frac{1}{nr}\sum_{k\in N}\ln\lambda_k \\ &= \sum_{k\in N_j}\mu_{jk}^* + \frac{1}{2}r\sigma^2\sum_{k\in N_j\backslash\overline{N}_i}\frac{1}{\left(d_k+1\right)^2} + \frac{1}{r}\ln\lambda_j\end{aligned}$$

implying that μ^* also solves the system of equations (21). Hence, t^* is Pareto efficient.

$\square$

Lemma 7. *A linear profile of transfer rules $t = (\alpha, \beta, \mu)$ is Pareto efficient if $\forall ij$ s.t. $G_{ij} = 1$,*

$$\begin{cases}\alpha_{ij} &= \frac{1}{2}\left(1 - \sum_{k\in N_i\backslash\{j\}}\alpha_{ik} + \sum_{k\in N_{ij}}\beta_{jki} + \gamma_{ij}\right) \\ \beta_{ijk} &= \frac{1}{2}\Big[\alpha_{ki} - \alpha_{kj} + \sum_{h\in N_{ijk}}\left(\beta_{ihk} - \beta_{jhk}\right) \\ &\quad - \sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ihk} + \sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh} + \gamma_{ij}\Big] \quad \forall k\in N_{ij} \\ \gamma_{ij} &= \frac{\rho}{1+(d_{ij}+1)\rho}\Big[\sum_{k\in N_i\backslash\overline{N}_j}\left(\alpha_{ki} - \sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right) \\ &\quad - \sum_{k\in N_j\backslash\overline{N}_i}\left(\alpha_{kj} - \sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right) \\ &\quad - \sum_{k\in N_{ij}}\left(\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh} - \sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\Big]\end{cases} \tag{25}$$

Proof of Proposition 4

Proof. We begin by proving the first part, which only involves triads. We rewrite (12) in the following way:

$$\begin{cases} 2\alpha_{ij} + \sum_{k\in N_i\setminus\{j\}} \alpha_{ik} - \gamma_{ij} = 1, \quad \forall G_{ij} = 1 & (1) \\ \alpha_{ki} - \alpha_{kj} + \gamma_{ij} = 0 \quad \forall k \in N_{ij}, \quad \forall G_{ij} = 1; & (2) \\ \gamma_{ij} = \frac{\rho}{1+(d_{ij}+1)\rho}\left(\sum_{k\in N_i\setminus\overline{N}_j} \alpha_{ki} - \sum_{k\in N_j\setminus\overline{N}_i} \alpha_{kj}\right), \quad \forall G_{ij} = 1; & (3) \end{cases}$$

In matrix form we write

$$\begin{bmatrix} \tilde{A} \\ M \end{bmatrix} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \begin{pmatrix} \tilde{b} \\ \mathbf{0} \end{pmatrix} \qquad \begin{matrix} (1) \wedge (3) \\ (2) \end{matrix}$$

where α, γ are both $\sum_i d_i$-dimensional vectors, $\tilde{A}$ is a $(2\sum_i d_i) \times (2\sum_i d_i)$ square matrix, $\tilde{b} := \begin{pmatrix} \mathbf{1}_{\sum_i d_i} \\ \mathbf{0}_{\sum_i d_i} \end{pmatrix}$ is a $(2\sum_i d_i)$-dimensional vector, M is a $\left(\sum_{G_{ij}=1} d_{ij}\right) \times (2\sum_i d_i)$ rectangular matrix, and $\mathbf{0}$ is a $\left(\sum_{G_{ij}=1} d_{ij}\right)$-dimensional vector. The upper block $\tilde{A}\begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \tilde{b}$ corresponds to equations in (1) and (3), while the lower block $M\begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = \mathbf{0}$ corresponds to equations in (2).

If the above system admits a solution (α, γ), then the solution must be unique. To see this, notice that the consumption plan induced by α is given by $x_i(e) = \left(1 - \sum_{j\in N_i} \alpha_{ij}\right) e_i + \sum_{j\in N_i} \alpha_{ji}$, so two distinct α will lead to two distinct consumption plans. By Proposition 2, the Pareto efficient consumption plan is unique, so the Pareto efficient α must also be unique. As γ is defined as a function α, γ must also be unique. Hence, if the above system admits a solution, the solution must be unique, and the matrix $\tilde{A}$ must have full rank.

We now show that the equations in (2) cannot be linearly independent from those in (1) and (3); in other words, for each equation in (2), there exists a nonzero vector $\xi \in \mathbb{R}^{2\sum_i d_i + \sum_{G_{ij}=1} d_{ij}}$ such that its entry corresponding to that equation in (2) is nonzero and that

$$\xi' \begin{bmatrix} \tilde{A} & \tilde{b} \\ M & \mathbf{0} \end{bmatrix} = (0, 0, ..., 0)_{2\sum_i d_i + 1}.$$

For any linked pair ij, multiply $③_{ij}$ (the ij-th equation in ③) with $(1 + (d_{ij} + 1)\rho)$, obtaining $[1 + (d_{ij} + 1)\rho]\gamma_{ij} = \rho \left(\sum_{h\in N_i \setminus \overline{N}_j} \alpha_{hi} - \sum_{h\in N_j \setminus \overline{N}_i} \alpha_{hj}\right)$, which is equivalent to

$$[1 + (d_{ij} + 1)\rho]\gamma_{ij} = \rho \left[\sum_{h\in N_i} \alpha_{hi} - \sum_{h\in N_j} \alpha_{hj} - \sum_{h\in N_{ij}} (\alpha_{hi} - \alpha_{hj}) - \alpha_{ji} + \alpha_{ij}\right] \quad ④_{ij}.$$

Adding $②_{ijh}$ for all $h \in N_{ij}$ to $④_{ij}$, we get

$$[1 + (d_{ij} + 1)\rho]\gamma_{ij} = \rho \left[\sum_{h\in N_i} \alpha_{hi} - \sum_{h\in N_j} \alpha_{hj} + d_{ij}\gamma_{ij} - \alpha_{ji} + \alpha_{ij}\right]$$

which is equivalent to $⑤_{ij}$: $(1 + \rho)\gamma_{ij} = \rho \left(\sum_{h\in N_i} \alpha_{hi} - \sum_{h\in N_j} \alpha_{hj} + \alpha_{ij} - \alpha_{ji}\right)$. Given any linked triads ijk, summing up $⑤_{ij}$, $⑤_{jk}$, $⑤_{ki}$, we have

$$(1 + \rho)(\gamma_{ij} + \gamma_{jk} + \gamma_{ki}) = \rho(\alpha_{ij} + \alpha_{jk} + \alpha_{ki} - \alpha_{ji} - \alpha_{kj} - \alpha_{ik}) \quad ⑥_{ijk}.$$

Moreover, by $②_{ijk}$, $②_{jki}$, $②_{kij}$, we have

$$\alpha_{ij} - \alpha_{ik} + \alpha_{jk} - \alpha_{ji} + \alpha_{ki} - \alpha_{kj} = -\gamma_{jk} - \gamma_{ki} - \gamma_{ij} \quad ⑦_{ijk}.$$

Hence, ⑥ $- \rho \times$ ⑦ gives $(1 + 2\rho)(\gamma_{ij} + \gamma_{jk} + \gamma_{ki}) = 0$. For $n = 3$ and $\rho > -\frac{1}{2}$, or for $n \geq 4$, we have $1 + 2\rho > 0$ and thus

$$\gamma_{ij} + \gamma_{jk} + \gamma_{ki} = 0. \quad ⑧_{ijk}$$

Taking $①_{ki} - ①_{kj} + ②_{ijk}$, we obtain $\gamma_{ij} - \gamma_{kj} + \gamma_{ki} = 0$. By $③_{jk} + ③_{kj}$, we have $\gamma_{jk} + \gamma_{kj} = 0$ and thus

$$\gamma_{ij} + \gamma_{jk} + \gamma_{ki} = 0. \quad ⑨_{ijk}$$

Then $⑧_{ijk} - ⑨_{ijk}$ leads to $0 = 0$. But notice that the vector $\xi \in \mathbb{R}^{2\sum_i d_i + \sum_{G_{ij}=1} d_{ij}}$

that characterizes all the row operations conducted above must be nonzero. In particular, we must have $\xi_{①_{ki}} \neq 0$, $\xi_{①_{kj}} \neq 0$, because $①_{ki}, ①_{kj}$ are used to obtain $⑨_{ijk}$ and nowhere else. As $\tilde{A}$ is of full row rank, this leads to the conclusion that ② must be a linear combination of ① and ③.

We now prove the second part, the statement for cycles of any size. Note that we still have $⑤_{ij}:\ (1+\rho)\,\gamma_{ij} = \rho\left(\sum_{h\in N_i}\alpha_{hi} - \sum_{h\in N_j}\alpha_{hj} + \alpha_{ij} - \alpha_{ji}\right)$. Given any cycle $i_1 i_2 ... i_m i_1$, summing up $⑤_{i_1 i_2}, ⑤_{i_2 i_3}, ..., ⑤_{i_m i_1}$, we have

$$(1+\rho)\left(\gamma_{i_1 i_2} + \gamma_{i_2 i_3} + ... + \gamma_{i_m i_1}\right) = \rho\left(\alpha_{i_1 i_2} + ... + \alpha_{i_m i_1} - \alpha_{i_2 i_1} - ... - \alpha_{i_1 i_m}\right) \quad ⑩$$

By $①_{i_1 i_2} - ①_{i_2 i_1}$ and $\gamma_{ij} + \gamma_{ji} = 0$, we have $\alpha_{i_1 i_2} - \alpha_{i_2 i_1} = \sum_{h\in N_{i_2}}\alpha_{i_2 h} - \sum_{h\in N_{i_1}}\alpha_{i_1 h} + 2\gamma_{i_1 i_2}$. Summing over $i_1 i_2, ..., i_m i_1$,

$$\alpha_{i_1 i_2} + ... + \alpha_{i_m i_1} - \alpha_{i_2 i_1} - ... - \alpha_{i_1 i_m} = 2\left(\gamma_{i_1 i_2} + \gamma_{i_2 i_3} + ... + \gamma_{i_m i_1}\right) \quad ⑪$$

Then $⑩ + \rho \times ⑪$ gives $(1-\rho)\left(\gamma_{i_1 i_2} + \gamma_{i_2 i_3} + ... + \gamma_{i_m i_1}\right) = 0$. For $\rho < 1$, we have $\gamma_{i_1 i_2} + \gamma_{i_2 i_3} + ... + \gamma_{i_m i_1} = 0$. □

Proof of Proposition 5

Proof. Write system (12) in the following form:

$$\begin{cases} 2\alpha_{ij} + \sum_{k\in N_i\setminus\{j\}}\alpha_{ik} - \gamma_{ij} = 1, \quad \forall G_{ij} = 1 & ① \\ \gamma_{ij} = \frac{\rho}{1+(d_{ij}+1)\rho}\left(\sum_{k\in N_i\setminus\overline{N}_j}\alpha_{ki} - \sum_{k\in N_j\setminus\overline{N}_i}\alpha_{kj}\right), \quad \forall G_{ij} = 1; & ② \end{cases}$$

This is a system of $2\sum_i d_i$ equations in $2\sum_i d_i$ variables (α, γ). Notice that this system can have at most one solution by Proposition 2, as each distinct solution to the above system will define a distinct consumption plan.

Write system (14) in the following form: $\forall ij$ s.t. $G_{ij} = 1$, and $\forall i \in N$

$$\begin{cases} \alpha_{ji} = \Lambda_j - \frac{\rho}{1-\rho}\left(\sum_{k\in N_i}\alpha_{ki} + \alpha_{ii}\right) \quad \forall G_{ij} = 1 & ③_{ji} \\ \alpha_{ii} = \Lambda_i - \frac{\rho}{1-\rho}\left(\sum_{k\in N_i}\alpha_{ki} + \alpha_{ii}\right) \quad \forall i \in N & ③_{ii} \\ \alpha_{ii} + \sum_{k\in N_i}\alpha_{ik} = 1 \quad \forall i \in N & ④_i \end{cases}$$

This is a system of $(\sum_i d_i + 2n)$ equations in $(\sum_i d_i + 2n)$ variables (α, Λ). Suppose that this system has a unique solution. [34]

Notice, as in the proof of Proposition 4, (1) and (2) imply that

$$(1+\rho)\,\gamma_{ij} = \rho\left(\sum_{h\in N_i}\alpha_{hi} - \sum_{h\in N_j}\alpha_{hj} + \alpha_{ij} - \alpha_{ji}\right) \quad (5)_{ij}.$$

By $(3)_{ji} - (3)_{ij}$ we have $\alpha_{ji} - \alpha_{ij} = \Lambda_j - \Lambda_i - \frac{\rho}{1-\rho}\left(\sum_{k\in N_i}\alpha_{ki} + \alpha_{ii} - \sum_{k\in N_j}\alpha_{kj} - \alpha_{jj}\right)$, which implies that

$$\rho\left(\sum_{k\in N_i}\alpha_{ki} - \sum_{k\in N_j}\alpha_{kj}\right) = (1-\rho)\left(\Lambda_j - \Lambda_i + \alpha_{ij} - \alpha_{ji}\right) - \rho\left(\alpha_{ii} - \alpha_{jj}\right),$$

leading to

$$\rho\left(\sum_{k\in N_i}\alpha_{ki} - \sum_{k\in N_j}\alpha_{kj} + \alpha_{ij} - \alpha_{ji}\right) = (1-\rho)\left(\Lambda_j - \Lambda_i\right) + \alpha_{ij} - \alpha_{ji} - \rho\left(\alpha_{ii} - \alpha_{jj}\right).$$

By $(5)_{ij}$, we have $(1+\rho)\,\gamma_{ij} = (1-\rho)\left(\Lambda_j - \Lambda_i\right) + \alpha_{ij} - \alpha_{ji} - \rho\left(\alpha_{ii} - \alpha_{jj}\right)$. Also, by $(1)_{ij}$ and $(4)_{ii}$, we have $\alpha_{ij} = \alpha_{ii} + \gamma_{ij}$. Plugging this into the above, we have

$$(1+\rho)\,\gamma_{ij} = (1-\rho)\left(\Lambda_j - \Lambda_i\right) + \gamma_{ij} - \gamma_{ji} + (1-\rho)\left(\alpha_{ii} - \alpha_{jj}\right)$$

Writing $\Lambda_{ij} := \Lambda_i - \Lambda_j$, we deduce $(1-\rho)\left(\gamma_{ij} - \Lambda_{ij} + \alpha_{ii} - \alpha_{jj}\right) = 0$, which implies

$$\Lambda_{ij} = \gamma_{ij} + \alpha_{ii} - \alpha_{jj}. (6)_{ij}$$

Now, by $(3)_{ii}$, we have $(1-\rho)\,\alpha_{ii} = (1-\rho)\,\Lambda_i - \rho\left(\sum_{k\in N_i}\alpha_{ki} + \alpha_{ii}\right)$.This implies that $\rho\left(\sum_{k\in N_i}\alpha_{ki}\right) = (1-\rho)\,\Lambda_i - \alpha_{ii}$ and thus

$$\rho\left(\sum_{k\in N_i}\alpha_{ki} - \sum_{k\in N_j}\alpha_{kj}\right) = (1-\rho)\,\Lambda_{ij} - \alpha_{ii} + \alpha_{jj}.$$

Again, by $(5)_{ij}$, we have $(1+\rho)\,\gamma_{ij} = (1-\rho)\,\Lambda_{ij} - \alpha_{ii} + \alpha_{jj} + \rho\left(\alpha_{ij} - \alpha_{ji}\right)$. Then, by

[34] It indeed has a unique solution given by Proposition 6.

$\alpha_{ij} = \alpha_{ii} + \gamma_{ij}$, a result of $①_{ij}$ and $④_{ij}$, we have

$$(1+\rho)\,\gamma_{ij} = (1-\rho)\,\Lambda_{ij} - \alpha_{ii} + \alpha_{jj} + \rho\,(\alpha_{ii} - \alpha_{jj} + 2\gamma_{ij})$$

This implies that $(1-\rho)\,\gamma_{ij} = (1-\rho)\,(\Lambda_{ij} - \alpha_{ii} + \alpha_{jj})$, leading to

$$\Lambda_{ij} = \gamma_{ij} + \alpha_{ii} - \alpha_{jj}. ⑦_{ij}$$

Notice that $⑥_{ij}$ - $⑦_{ij}$ gives $0 = 0$. Given the previous row operations, we are sure that equations $③_{ij}$, $③_{ji}$ and $③_{ii}$ are all used in the process and are not canceled out. As ③ and ④ have full rank, equations in ③ must be linearly independent. Hence, we know that each equation of ③ must be linearly dependent on ①, ② and ④. $\square$

Proof of Proposition 6

Proof. Let $\overline{G} := G + \mathbf{I}_n$ so that $\overline{G}_{ii} = 1\ \forall i \in N$. The optimality conditions given in equation (14.1) and (14.2) can be rewritten as

$$\alpha_{ji} = \overline{G}_{ij}\left(\Lambda_j - \frac{\rho}{1-\rho}\sum_{k\in N}\overline{G}_{ik}\alpha_{ki}\right) \tag{26}$$

Let $\overline{\alpha}_i := (\alpha_{1i}, \alpha_{2i}, \ldots, \alpha_{ni})'$ denote the vector of i's inflow shares, $\mathbf{\Lambda} = (\Lambda_1, \Lambda_2, \ldots, \Lambda_n)'$ the vector of rescaled constraint multipliers, and g_i represent the i-th column of $\overline{G}$. Then (26) can be rewritten in vector form as

$$\left(\mathbf{I} + \frac{\rho}{1-\rho} g_i g_i^{'}\right)\alpha_i = diag\,(g_i)\,\mathbf{\Lambda}$$

where $diag\,(g_i)$ is a diagonal matrix with g_i's entries on the diagonal. Left-multiplying both sides by $\left(\mathbf{I} - \frac{\rho}{1+\rho d_i} g_i g_i^{'}\right)$, which is well-defined for any $\rho > -\frac{1}{n-1}$ and any G, we have

$$\alpha_i = \left(\mathbf{I} - \frac{\rho}{1+\rho d_i} g_i g_i^{'}\right) diag\,(g_i)\,\mathbf{\Lambda}$$

As $g_i g_i^{'} \cdot diag\left(g_i\right) = g_i g_i^{'}$, the above becomes

$$\alpha_i = \left(diag\left(g_i\right) - \frac{\rho}{1+\rho d_i} g_i g_i^{'} \right) \mathbf{\Lambda} \tag{27}$$

Now, notice that (14.3) implies

$$1 = \sum_{j \in N} \alpha_{ij} = \left(d_i + 1\right) \Lambda_i - \sum_{j \in N} \overline{G}_{ij} \left(\frac{\rho}{1+\rho d_j} \sum_{k} \overline{G}_{jk} \Lambda_k \right) \tag{28}$$

and thus we have

$$\Lambda_i = \frac{1}{d_i + 1} \left(1 + \sum_{j \in \overline{N}_i} \sum_{k \in \overline{N}_j} \frac{\rho}{1+\rho d_j} \Lambda_k \right).$$

This establishes the recursive representation of the solution.

To obtain the closed-form solution, rewrite equation (28) as

$$\mathbf{1} = \sum_{i \in N} \left(diag\left(g_i\right) - \frac{\rho}{1+\rho d_i} g_i g_i^{'} \right) \mathbf{\Lambda} = \left(\overline{D} - \overline{G} \Psi \overline{G}\right) \Lambda$$

where $\overline{D}$ is a diagonal matrix with its i-th diagonal entry being $d_i + 1$, and Ψ is a diagonal matrix with its i-th diagonal entry being $\frac{\rho}{1+\rho d_i}$. Notice that $\forall \xi \in \mathbb{R}^n \backslash \{\mathbf{0}\}$,

$$\begin{aligned}
\xi^{'} \left(\overline{D} - \overline{G} \Psi \overline{G}\right) \xi &= \sum_{i \in N} \left(d_i + 1\right) \xi_i^2 - \sum_{i \in N} \frac{\rho}{1+\rho d_i} \left(\sum_{j \in \overline{N}_i} \xi_j \right)^2 \\
&\geq \sum_{i \in N} \left(d_i + 1\right) \xi_i^2 - \sum_{i \in N} \frac{1}{1+d_i} \left(\sum_{j \in \overline{N}_i} \xi_j \right)^2 \\
&\geq \sum_{i \in N} \left(d_i + 1\right) \xi_i^2 - \sum_{i \in N} \frac{1}{1+d_i} \cdot \left(1 + d_i\right) \sum_{j \in \overline{N}_i} \xi_j^2 \\
&= \sum_{i \in N} \left(d_i + 1\right) \xi_i^2 - \sum_{i \in N} \left(d_i + 1\right) \xi_i^2 \\
&= 0
\end{aligned}$$

where the equality holds if and only if $\rho = 1$ and $\xi = c \cdot \mathbf{1}$ for some $c > 0$. Hence,

$\forall\rho \in \left(-\frac{1}{n-1},1\right)$, $\left(\overline{D}-\overline{G}\Psi\overline{G}\right)$ is positive definite and thus invertible. Hence,

$$\begin{aligned}\Lambda &= \left(\overline{D}-\overline{G}\Psi\overline{G}\right)^{-1}\mathbf{1},\\ \alpha_i &= \left(diag\left(g_i\right)-\frac{\rho}{1+\rho d_i}g_i g_i^{'}\right)\left(\overline{D}-\overline{G}\Psi\overline{G}\right)^{-1}\mathbf{1}.\end{aligned}$$

Finally, we solve for the inverse matrix above as a series of powers of $\overline{G}$. Notice that

$$\left(\overline{D}-\overline{G}\Psi\overline{G}\right)^{-1}=\left(\overline{D}^{\frac{1}{2}}\left(\mathbf{I}-\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)\overline{D}^{\frac{1}{2}}\right)^{-1}=\overline{D}^{-\frac{1}{2}}\left(\mathbf{I}-\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)^{-1}\overline{D}^{-\frac{1}{2}}$$

where the middle term $\left(\mathbf{I}-\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)$ is also invertible and positive definite for $\rho \in \left(-\frac{1}{n-1},1\right)$ due to the positive definiteness of $\overline{D}-\overline{G}\mathbf{\Psi}\overline{G}$ and the invertibility of $\overline{D}$. For $\rho \in (0,1)$, notice that $\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}$ is also positive definite, so its eigenvalues must be positive. Also, its largest eigenvalue φ_{max} must be smaller than 1. Otherwise, there exists a nonzero vector ξ such that

$$\xi^{'}\left(I-\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{\frac{1}{2}}\right)\xi=\left(1-\varphi_{max}\right)\xi^{'}\xi<0$$

contradicting the positive definiteness of $\left(\mathbf{I}-\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)$. Then, we may write

$$\left(\mathbf{I}-\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)^{-1}=\mathbf{I}+\sum_{k=1}^{\infty}\left(\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)^{k}$$

and thus

$$\begin{aligned}\left(\overline{D}-\overline{G}\Psi\overline{G}\right)^{-1}&=\overline{D}^{-1}+\overline{D}^{-\frac{1}{2}}\sum_{k=1}^{\infty}\left(\overline{D}^{-\frac{1}{2}}\overline{G}\Psi\overline{GD}^{-\frac{1}{2}}\right)^{k}\overline{D}^{-\frac{1}{2}}\\ &=\overline{D}^{-1}+\sum_{k=1}^{\infty}\left(\overline{D}^{-1}\overline{Q}\right)^{k}\overline{D}^{-1}\end{aligned}$$

where $\overline{Q}:=\overline{G}\Psi\overline{G}$ can be interpreted as the weighted square of the extended adjacency matrix. Consider the set of all paths of length q between i and j under G as

$$\Pi_{ij}^{q}\left(G\right)=\left\{\left(i_0,i_1,i_2,\ldots i_q\right)\mid i_0=i,\ i_q=j\ and\ \overline{G}_{i_n i_{n+1}}=1\ for\ n=0,1,\ldots q-1\right\}$$

For every $\pi_{ij} \in \Pi_{ij}^{q}(G)$, let $W(\pi_{ij})$ denote the weight associated to this path. It is not difficult to see that,

$$W(\pi_{ij}) = \frac{1}{d_i+1}\frac{\rho}{1+\rho d_{i_1}}\frac{1}{d_{i_2}+1}\frac{\rho}{1+\rho d_{i_3}}\cdots\frac{1}{d_j+1}$$

Then

$$\begin{aligned}
\Lambda_i &= \left[\left(\overline{D}-\overline{G}\Psi\overline{G}\right)^{-1}\mathbf{1}\right]_i \\
&= \left(\overline{D}^{-1}\mathbf{1}\right)_i + \left(\sum_{k=1}^{\infty}\left(\overline{D}^{-1}\overline{Q}\right)^k\overline{D}^{-1}\mathbf{1}\right)_i \\
&= \frac{1}{d_i+1} + \left[\sum_{k=1}^{\infty}\left(\overline{D}^{-1}\overline{G}\Psi\overline{G}\right)^k\right]_i\overline{D}^{-1}\mathbf{1} \\
&= \frac{1}{d_i+1} + \sum_{j\in N}\left[\sum_{k=1}^{\infty}\left(\overline{D}^{-1}\overline{G}\Psi\overline{G}\right)^k\right]_{ij}\cdot\frac{1}{d_j+1} \\
&= \frac{1}{d_i+1} + \sum_{j\in N}\sum_{k=1}^{\infty}\left(\overline{D}_{[i]}^{-1}\overline{G}\Psi\overline{G}\overline{D}^{-1}...\overline{D}^{-1}\overline{G}\Psi\overline{G}_{(j)}\right)\frac{1}{d_j+1} \\
&= \frac{1}{d_i+1} + \sum_{j\in N}\sum_{q=1}^{\infty}\sum_{\pi_{ij}\in\Pi_{ij}^{2q}}\left(\frac{1}{d_i+1}\cdot\frac{\rho}{1+\rho d_{i_1}}\cdot\frac{1}{d_{i_2}+1}\cdot...\right)\frac{1}{d_j+1} \\
&= \frac{1}{d_i+1} + \sum_{q=1}^{\infty}\sum_{j\in N}\sum_{\pi_{ij}\in\Pi_{ij}^{2q}}W(\pi_{ij})
\end{aligned}$$

This concludes the proof. □

Proof of Proposition 7

Proof. (i) is immediate from (12.2). (ii) follows from differencing (15). (iii) follows from $(7)_{ij}$ in the proof of Proposition 5. □

Supplement to "Informal Risk Sharing with Local Information"

Attila Ambrus[35], Wayne Yuan Gao[36], Pau Milán[37]

Appendix B: Supplementary Materials

B.1 Proofs of Lemmas

Lemma 1. *$\mathcal{T}^*$ with $\langle\cdot,\cdot\rangle$ forms a Hilbert space.*

Proof. We first show that $\langle\cdot,\cdot\rangle$ is a well-defined inner product. Symmetry immediately follows from the definition. Linearity in the first argument follows from the linearity of the expectation operator:

$$\begin{aligned}\langle \alpha s+\beta t,\ r\rangle &= \mathbb{E}\left[\sum_{G_{ij}=1}\left(\alpha s_{ij}+\beta t_{ij}\right)r_{ij}\right]=\alpha\mathbb{E}\left[\sum_{G_{ij}=1}s_{ij}r_{ij}\right]+\beta\mathbb{E}\left[\sum_{G_{ij}=1}t_{ij}r_{ij}\right]\\ &=\alpha\langle s,\ r\rangle+\beta\langle t,\ r\rangle .\end{aligned}$$

Positive definiteness is also obvious: $\langle t,t\rangle=\mathbb{E}\left[\sum_{G_{ij}=1}t_{ij}^2(\omega)\right]\geq 0$ and $\langle t,t\rangle=0$ if and only if $t=\mathbf{0}$, i.e., $t_{ij}(\omega)=0$ for all linked ij and $\mathbb{P}$-almost all $\omega\in\Omega$.

We then show that $\mathcal{T}^*$ is a linear space. $\forall s,t\in\mathcal{T}^*$, $\forall\alpha,\beta\in\mathbb{R}$, $\alpha s(I_{ij})+\beta t(I_{ij})$ is also $\sigma(I_{ij})$-measurable, and

$$\alpha s_{ij}(\omega)+\beta t_{ij}(\omega)=-\left(\alpha s_{ji}(\omega)+\beta t_{ji}(\omega)\right).$$

Hence, $\alpha s+\beta t\in\mathcal{T}$. Furthermore,

$$\langle\alpha s+\beta t,\alpha s+\beta t\rangle=\alpha^2\langle s,s\rangle+2\alpha\beta\langle s,t\rangle+\beta^2\langle t,t\rangle<\infty\quad ,$$

[35] Ambrus: Department of Economics, Duke University, Durham, NC 27708, USA, attila.ambrus@duke.edu.

[36] Gao: Department of Economics, Yale University, New Haven, CT 06511, USA, wayne.gao@yale.edu.

[37] Miln: Department of Economics, Universitat Autnoma de Barcelona, Edifici B, Bellaterra 08193, Spain, pau.milan@uab.cat.

because

$$\langle s,t\rangle^2 = \left(\sum_{ij}\mathbb{E}\left[s_{ij}t_{ij}\right]\right)^2 \leq \left(\sum_{ij}\left(\mathbb{E}\left[s_{ij}^2\right]\mathbb{E}\left[t_{ij}^2\right]\right)^{\frac{1}{2}}\right)^2$$
$$\leq \sum_{ij}\mathbb{E}\left[s_{ij}^2\right]\cdot\sum_{ij}\mathbb{E}\left[t_{ij}^2\right] = \langle s,s\rangle\cdot\langle t,t\rangle < \infty.$$

by applying the Cauchy-Schwarz inequality twice. So $\alpha s+\beta t\in\mathcal{T}^*$.

We finally show that $\mathcal{T}^*$ with $\langle\cdot,\cdot\rangle$ is complete. Taking $\left\{t^{(n)}\right\}$ to be a Cauchy sequence in $\mathcal{T}^*$, then$\left\{t_{ij}^{(n)}\right\}$ must be a Cauchy sequence in a $L^2(\mathbb{P})$ space with the inner product

$$\langle s_{ij},t_{ij}\rangle_{ij} := \mathbb{E}\left[s_{ij}(\omega)\,t_{ij}(\omega)\right].$$

As the $L^2(\mathbb{P})$ space is complete, $t_{ij}^{(n)}$ must converge to some t_{ij} in it with $\mathbb{E}\left[t_{ij}^2(\omega)\right]<\infty$. Define $t:=(t_{ij})_{G_{ij}=1}$. Clearly, $t\in\mathcal{T}^*$, and $t^{(n)}\to t$. Hence, $\mathcal{T}^*$ with $\langle\cdot,\cdot\rangle$ forms a Hilbert space. ◻

Lemma 2. *The objective function in (5)*

$$J(t) := \mathbb{E}\left[\sum_{k\in N}\lambda_k u_k\left(e_k-\sum_{h\in N_k}t_{kh}\right)\right]$$

is concave on $\mathcal{T}^*$.

Proof. $\forall s,t\in\mathcal{T}^*$, $\forall\alpha\in[0,1]$,

$$\begin{aligned}
&J(\alpha s+(1-\alpha)t)\\
&=\mathbb{E}\left[\sum_i\lambda_i u_i\left(e_i-\sum_{j\in N_i}\left(\alpha s_{ij}(\omega)+(1-\alpha)t_{ij}(\omega)\right)\right)\right]\\
&=\sum_i\lambda_i\mathbb{E}\left[u_i\left(\alpha\left(e_i-\sum_{j\in N_i}s_{ij}(\omega)\right)+(1-\alpha)\left(e_i-\sum_{j\in N_i}t_{ij}(\omega)\right)\right)\right]\\
&\geq\sum_i\lambda_i\mathbb{E}\left[\alpha u_i\left(e_i-\sum_{j\in N_i}s_{ij}(\omega)\right)+(1-\alpha)u_i\left(e_i-\sum_{j\in N_i}t_{ij}(\omega)\right)\right]\\
&=\alpha\mathbb{E}\left[\sum_i\lambda_i u_i\left(e_i-\sum_{j\in N_i}s_{ij}(\omega)\right)\right]+(1-\alpha)\mathbb{E}\left[\sum_i\lambda_i u_i\left(e_i-\sum_{j\in N_i}t_{ij}(\omega)\right)\right]\\
&=\alpha J(s)+(1-\alpha)J(t).
\end{aligned}$$

◻

Lemma 3. *J is twice Fréchet-differentiable.*

Proof. $\forall s, t \in \mathcal{T}^*$, for $\alpha > 0$,

$$
\begin{aligned}
&\frac{J(t+\alpha s) - J(t)}{\alpha} \\
&= \mathbb{E}\left[\sum_i \lambda_i \left[\frac{u_i\left(e_i - \sum_{j\in N_i} t_{ij}(\omega) - \alpha \sum_{j\in N_i} s_{ij}(\omega)\right) - u_i\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right)}{\alpha}\right]\right] \\
&= \mathbb{E}\left[\sum_i \lambda_i \left[-\frac{u_i^{'}\left(e_i - \sum_{j\in N_i} t_{ij}(\omega) - \alpha \tilde{s}(\omega)\right) \cdot \alpha \sum_{j\in N_i} s_{ij}(\omega)}{\alpha}\right]\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{for some } \tilde{s}_{ij}(\omega) \text{ between 0 and } \sum_{j\in N_i} s_{ij}(\omega) \\
&= -\mathbb{E}\left[\sum_i \lambda_i \left[u_i^{'}\left(e_i - \sum_{j\in N_i} t_{ij}(\omega) - \alpha \tilde{s}\right) \cdot \sum_{j\in N_i} s_{ij}(\omega)\right]\right] \\
&\to -\mathbb{E}\left[\sum_i \lambda_i \left[u_i^{'}\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right) \cdot \sum_{j\in N_i} s_{ij}(\omega)\right]\right] \text{ as } \alpha \to 0 \\
&= -\sum_i \lambda_i \mathbb{E}\left[u_i^{'}\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right) \mathbf{1}_{i\times N_i} \cdot s(\omega)\right] \\
&= \sum_i \lambda_i < f_i, s >
\end{aligned}
$$

where

$$
f_i(\omega) := -u_i^{'}\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right) \mathbf{1}_{i\times N_i}
$$

and $\mathbf{1}_{i\times N_i}$ is vector of 0 and 1s that equals 1 for the (directed) link ij for any $j \in N_i$ so that $\mathbf{1}_{i\times N_i} \cdot s(\omega) = \sum_{j\in N_i} s_{ij}(\omega)$. Define $J^{'}(t) : \mathcal{T}^* \to \mathbb{R}$ by

$$
J^{'}(t)\, s = \sum_i \lambda_i < f_i, s > .
$$

Clearly $J^{'}(t)$ is a linear operator on $\mathcal{T}^*$, and is thus the first-order Fréchet-derivative of J.

Similarly, $\forall t, v, w \in \mathcal{T}^*$,

$$
\begin{aligned}
&\frac{J'(t+\alpha w)v - J'(t)v}{\alpha} \\
&= \frac{\sum_i \lambda_i \mathbb{E}\left[u_i'\left(e_i - \sum_{j\in N_i} t_{ij}(\omega) - \alpha \sum_{j\in N_j} w_{ij}(\omega)\right)\cdot \sum_{j\in N_i} v_{ij}(\omega)\right]}{\alpha} \\
&\quad - \frac{\sum_i \lambda_i \mathbb{E}\left[u_i'\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right)\cdot \sum_{j\in N_i} v_{ij}(\omega)\right]}{\alpha} \\
&= \mathbb{E}\left[\sum_i \lambda_i \left[\frac{u_i''\left(e_i - \sum_{j\in N_i} t_{ij}(\omega) - \alpha\tilde{w}(\omega)\right)\cdot \alpha \sum_{j\in N_i} w_{ij}(\omega)}{\alpha} \sum_{j\in N_i} v_{ij}(\omega)\right]\right] \\
&\qquad\qquad\qquad\qquad \text{for some } \tilde{w}(\omega) \text{ between 0 and } \sum_{j\in N_i} w_{ij}(\omega) \\
&= \mathbb{E}\left[\sum_i \lambda_i \left[u_i''\left(e_i - \sum_{j\in N_i} t_{ij}(\omega) - \alpha\tilde{w}(\omega)\right)\sum_{j\in N_i} w_{ij}(\omega)\cdot \sum_{j\in N_i} v_{ij}(\omega)\right]\right] \\
&\to \mathbb{E}\left[\sum_i \lambda_i \left[u_i''\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right)\sum_{j\in N_i} w_{ij}(\omega)\cdot \sum_{j\in N_i} v_{ij}(\omega)\right]\right] \text{ as } \alpha \to 0 \\
&= \sum_i \lambda_i E\left[u_i''\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right)\left(\mathbf{1}_{i\times N_i}\cdot w(\omega)\right)\left(\mathbf{1}_{i\times N_i}\cdot v(\omega)\right)\right]
\end{aligned}
$$

which is clearly bilinear in v and w, so J is twice Fréchet-differentiable. □

Lemma 4. *For any $t \in \mathcal{T}^*$ that solves (6), we have*

$$J'(t) = \mathbf{0}.$$

Proof. To solve (6)

$$\max_{\tilde{t}_{ij}\in\mathbb{R}} J^{(ij,I_{ij})}\left(\tilde{t}_{ij}\right) := \mathbb{E}\left[\lambda_i u_i\left(e_i - \tilde{t}_{ij} - \sum_{h\in N_i} t_{ih}\right) + \lambda_j u_j\left(e_j + \tilde{t}_{ij} - \sum_{h\in j} t_{jh}\right)\middle| I_{ij}\right]$$

we first notice the objective function $J^{(ij,I_{ij})}\left(\tilde{t}_{ij}\right)$ is strictly concave in $\tilde{t}_{ij}$ on $\mathbb{R}$. Hence, the sufficient and necessary condition for optimality is given by the FOC:

$$\mathbb{E}\left[\lambda_i u_i'\left(e_i - \sum_{h\in N_i} t_{ih}(\omega)\right)\middle| I_{ij}\right] = \mathbb{E}\left[\lambda_j u_j'\left(e_j - \sum_{h\in N_j} t_{jh}(\omega)\right)\middle| I_{ij}\right]$$

Then, $\forall s \in \mathcal{T}^*$,

$$\begin{aligned}
J^{'}(t)\, s &= -\mathbb{E}\left[\sum_i \lambda_i \left[u_i^{'}\left(e_i - \sum_{j\in N_i} t_{ij}(\omega)\right)\cdot \sum_{j\in N_i} s_{ij}(\omega)\right]\right] \\
&= -\frac{1}{2}\sum_{G_{ij}=1}\mathbb{E}\left[\left(\lambda_i u_i^{'}\left(e_i - \sum_{h\in N_i} t_{ih}(\omega)\right) - \lambda_j u_j^{'}\left(e_j - \sum_{h\in N_j} t_{jh}(\omega)\right)\right)\cdot s_{ij}(\omega)\right] \\
&= -\frac{1}{2}\sum_i\sum_{j\in N_i}\mathbb{E}\left[s_{ij}(I_{ij})\cdot \mathbb{E}\left[\lambda_i u_i^{'}\left(e_i - \sum_{h\in N_i} t_{ih}(\omega)\right) - \lambda_j u_j^{'}\left(e_j - \sum_{h\in N_j} t_{jh}(\omega)\right)\middle| I_{ij}\right]\right] \\
&= -\frac{1}{2}\sum_i\sum_{j\in N_i}\mathbb{E}\left[s_{ij}(I_{ij})\cdot 0\right] \\
&= 0.
\end{aligned}$$

Hence $J^{'}(t) = \mathbf{0}$. □

Lemma 5. *The set of consumption plan induced by the profiles of transfer rules t in $\mathcal{T}^*$ is convex.*

Proof. Let $x, x^{'}$ be two profiles of consumption plans induced by $t, t^{'}$ respectively. Then $\forall\lambda \in [0,1]$,

$$\begin{aligned}
\lambda x_i(\omega) + (1-\lambda)\, x_i^{'}(\omega) &= \lambda\left[e_i - \sum_{j\in N_i} t_{ij}(\omega)\right] + (1-\lambda)\left[e_i - \sum_{j\in N_i} t_{ij}^{'}(\omega)\right] \\
&= e_i - \sum_{j\in N_u}\left[\lambda t_{ij}(\omega) + (1-\lambda)\, t_{ij}^{'}(\omega)\right]
\end{aligned}$$

Thus $\left(\lambda x + (1-\lambda)\, x^{'}\right)$ can be induced by $\left(\lambda t + (1-\lambda)\, t^{'}\right)$. $\mathcal{T}^*$, as a Hilbert space, is convex, so the set of consumption plans induced by the profiles of transfer rules in $\mathcal{T}^*$ must also be convex. □

Lemma 6. *Given any real vector $c \in \mathbb{R}^n$ such that $\sum_{i\in N} c_i = 0$, there exists a real vector $\mu \in \mathbb{R}^{\sum_i d_i}$ such that $\mu_{ik} + \mu_{ki} = 0$ for every linked pair ik and*

$$\sum_{k\in N_i} \mu_{ik} = c_i. \tag{29}$$

The solution is unique if and only if the network is minimally connected.

Proof. With the restrictions that $\mu_{ik} = -\mu_{ki}$ for all linked pair ik, (29) constitutes a system of n linear equations with $\frac{1}{2}\sum_{i\in N} d_i$ variables μ_{ik}. Summing up all the n equations, we have

$$0 = \sum_{i<k,\, G_{ik}=1} (\mu_{ik} + \mu_{ki}) = \sum_{i\in N} c_i = 0.$$

Hence, the n linear equations impose at most $(n-1)$ linearly independent conditions.

Viewing (29) in vector form,

$$C\mu = c$$

where C is a $n \times \frac{1}{2}\sum_{i\in N} d_i$ matrix. Note that in each column of C, denoted C_{ij} for $i < j$, there are either no nonzero entries (when $G_{ij} = 0$), or just two nonzero entries: 1 on the i-th row and -1 on the j-th row when $G_{ij} = 1$. Suppose $G_{ij} = 1$. Then, given any subset of individuals S that include i and j, if the rows of C corresponding to S are linearly dependent, these rows must sum to 0: this can be true only if all entries ik with $i \in S$ and $k \notin S$ are zero, implying that S form a component under G, and thus G is not connected if $\#(S) < n$. This is in contradiction with the supposition that G is connected when $\#(S) < n$. Hence, C must have exactly $(n-1)$ linearly independent rows.

Let $\tilde{C}$ and $\tilde{c}$ be the first $(n-1)$ rows of C and c. Then, as $\tilde{C}$ has full row rank, there always exists a solution to $\tilde{C}\mu = \tilde{c}$, and any of the solutions μ must also solve the equation $C\mu = c$. The solution is unique if and only if the component is minimally connected, when there are precisely $(n-1)$ links and thus $\tilde{C}$ is an invertible square matrix.

We can obtain one particular solution using the following algorithm. First, we can arbitrarily select a subset of links that minimally connect the nodes, i.e., the graph restricted to this subset of links is minimally connected. Then, there must exist at least one peripheral node, and we can first easily obtain μ_{ij} for all such peripheral nodes $i \in P_1 := \{k \in N : \ d_k = 1\}$. Then, we can look for new peripheral nodes ignoring the links involving nodes in P_1, and obtain μ_{ij} for all $i \in P_2 := \{k \in N : \ k \notin P_1 \wedge G_{kj} = 1 \text{ for some } j \in P_1\}$ with all previously calculated μ's taken as given. We iterate this process until we exhaust all nodes. Then we are left with a profile of μ that solves (29). $\square$

Lemma 7. *A linear profile of transfer rules $t = (\alpha, \beta, \mu)$ is Pareto efficient if $\forall ij$ s.t.*

$G_{ij} = 1$,

$$\begin{cases} \alpha_{ij} & = \frac{1}{2}\left(1 - \sum_{k\in N_i\backslash\{j\}} \alpha_{ik} + \sum_{k\in N_{ij}} \beta_{jki} + \gamma_{ij}\right) \\ \beta_{ijk} & = \frac{1}{2}\Big[\alpha_{ki} - \alpha_{kj} + \sum_{h\in N_{ijk}} (\beta_{ihk} - \beta_{jhk}) \\ & \quad - \sum_{h\in N_{ik}\backslash\overline{N}_j} \beta_{ihk} + \sum_{h\in N_{jk}\backslash\overline{N}_i} \beta_{jkh} + \gamma_{ij}\Big] \quad \forall k \in N_{ij} \\ \gamma_{ij} & = \frac{\rho}{1+(d_{ij}+1)\rho}\Big[\sum_{k\in N_i\backslash\overline{N}_j}\left(\alpha_{ki} - \sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right) \\ & \quad - \sum_{k\in N_j\backslash\overline{N}_i}\left(\alpha_{kj} - \sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right) \\ & \quad - \sum_{k\in N_{ij}}\left(\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh} - \sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\Big] \end{cases} \tag{25}$$

Proof. For each $k \in N_i\backslash\{j\}$, we then have

$$\begin{aligned} \sum_{k\in N_i\backslash\{j\}} t_{ik} &= e_i \sum_{k\in N_i\backslash\{j\}} \alpha_{ik} - \sum_{k\in N_i\backslash\{j\}} \alpha_{ki}e_k + \sum_{k\in N_i\backslash\{j\}}\sum_{h\in N_{ik}} \beta_{ikh}e_h + c_{ij} \\ &= e_i \sum_{k\in N_i\backslash\{j\}} \alpha_{ik} - \sum_{k\in N_{ij}} \alpha_{ki}e_k + \sum_{k\in N_{ij}}\left(\beta_{ikj}e_j + \sum_{h\in N_{ijk}}\beta_{ikh}e_h\right) + \sum_{k\in N_i\backslash\overline{N}_j}\sum_{h\in N_{ijk}}\beta_{ikh}e_h \\ &\quad - \sum_{k\in N_i\backslash\overline{N}_j}\alpha_{ki}e_k + \sum_{k\in N_{ij}}\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}e_h + \sum_{k\in N_i\backslash\overline{N}_j}\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}e_h + c_{ij} \end{aligned}$$

so that

$$\begin{aligned} t_{ij} &= \frac{1}{2}e_i - \frac{1}{2}e_j - \frac{1}{2}\sum_{k\in N_i\backslash\{j\}}\alpha_{ik}e_i + \frac{1}{2}\sum_{k\in N_j\backslash\{i\}}\alpha_{jk}e_j - \frac{1}{2}\sum_{k\in N_{ij}}(\beta_{ikj}e_j - \beta_{jki}e_i) \\ &\quad + \frac{1}{2}\sum_{k\in N_{ij}}\left[(\alpha_{ki} - \alpha_{kj})\,e_k - \sum_{h\in N_{ijk}}(\beta_{ikh} - \beta_{jkh})\,e_h\right] \\ &\quad - \frac{1}{2}\sum_{k\in N_i\backslash\overline{N}_j}\sum_{h\in N_{ijk}}\beta_{ikh}e_h + \frac{1}{2}\sum_{k\in N_j\backslash\overline{N}_i}\sum_{h\in N_{ijk}}\beta_{jkh}e_h \\ &\quad - \frac{1}{2}\frac{\rho}{1+(d_{ij}+1)\,\rho}\left(e_i + e_j + \sum_{k\in N_{ij}}e_k\right)\sum_{k\in N_{ij}}\left(\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh} - \sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right) \\ &\quad + \frac{1}{2}\frac{\rho}{1+(d_{ij}+1)\,\rho}\left(e_i + e_j + \sum_{k\in N_{ij}}e_k\right)\sum_{k\in N_i\backslash\overline{N}_j}\left(\alpha_{ki} - \sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right) \end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{2}\frac{\rho}{1+(d_{ij}+1)\rho}\left(e_i+e_j+\sum_{k\in N_{ij}}e_k\right)\sum_{k\in N_j\backslash\overline{N}_i}\left(\alpha_{kj}-\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\\
=&\frac{1}{2}\Bigg\{1-\sum_{k\in N_i\backslash\{j\}}\alpha_{ik}+\sum_{k\in N_{ij}}\beta_{jki}+\frac{\rho}{1+(d_{ij}+1)\rho}\Bigg[\sum_{k\in N_i\backslash\overline{N}_j}\left(\alpha_{ki}-\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right)\\
&-\sum_{k\in N_j\backslash\overline{N}_i}\left(\alpha_{kj}-\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)-\sum_{k\in N_{ij}}\left(\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}-\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\Bigg]\Bigg\}\cdot e_i\\
&-\frac{1}{2}\Bigg\{1-\sum_{k\in N_j\backslash\{i\}}\alpha_{jk}+\sum_{k\in N_{ij}}\beta_{ikj}+\frac{\rho}{1+(d_{ij}+1)\rho}\Bigg[\sum_{k\in N_j\backslash\overline{N}_i}\left(\alpha_{kj}-\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\\
&-\sum_{k\in N_i\backslash\overline{N}_j}\left(\alpha_{ki}-\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right)-\sum_{k\in N_{ij}}\left(\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}-\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right)\Bigg]\Bigg\}\cdot e_j\\
&+\frac{1}{2}\sum_{k\in N_{ij}}\Bigg\{\alpha_{ki}-\alpha_{kj}+\sum_{h\in N_{ijk}}(\beta_{ihk}-\beta_{jhk})-\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ihk}+\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\\
&+\frac{\rho}{1+(d_{ij}+1)\rho}\Bigg[\sum_{k\in N_i\backslash\overline{N}_j}\left(\alpha_{ki}-\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}\right)-\sum_{k\in N_j\backslash\overline{N}_i}\left(\alpha_{kj}-\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\\
&-\sum_{k\in N_{ij}}\left(\sum_{h\in N_{ik}\backslash\overline{N}_j}\beta_{ikh}-\sum_{h\in N_{jk}\backslash\overline{N}_i}\beta_{jkh}\right)\Bigg]\Bigg\}\cdot e_k+C_{ij}
\end{aligned}
$$

The last equality is obtained by collecting terms and switching summand indice. □

B.2 Uniqueness in Minimally Connected Networks

Proposition 10. *Under the independent CARA-Normal setting, if the network is minimally connected, then there is a unique profile of transfer rules in $\mathcal{T}^*$ that is Pareto efficient, and it takes the form of the local equal sharing rule.*

Proof. Consider minimally connected network G. For Pareto efficiency, we need for all linked pair ij

$$\frac{E_{ij}\left[u_i'(x_i)\right]}{E_{ij}\left[u_j'(x_j)\right]}=\alpha_{ij}.$$

As the network is minimally connected, we have $N_{ij} = \emptyset$. Notice that

$$\mathbb{E}\left[re^{-r\left(e_i - t_{ij} - \sum_{k\in N_i\backslash\{j\}} t_{ik}(e_i,e_k)\right)} \middle| e_i, e_j \right] = \alpha_{ij}\mathbb{E}\left[re^{-r\left(e_j + t_{ij} - \sum_{h\in N_j\backslash\{i\}} t_{jh}(e_j,e_h)\right)} \middle| e_i, e_j \right].$$

By independence,

$$\begin{aligned}
&\mathbb{E}\left[re^{-r\left(e_i - t_{ij} - \sum_{k\in N_i\backslash\{j\}} t_{ik}(e_i,e_k)\right)} \middle| e_i \right] = \alpha_{ij}\mathbb{E}\left[re^{-r\left(e_j + t_{ij} - \sum_{h\in N_j\backslash\{i\}} t_{jh}(e_j,e_h)\right)} \middle| e_j \right] \\
\Leftrightarrow & e^{-r(e_i - t_{ij})} \cdot \prod_{k\in N_i\backslash\{j\}} \mathbb{E}\left[e^{rt_{ik}(e_i,e_k)} \middle| e_i \right] = \alpha_{ij} e^{-r(e_j + t_{ij})} \cdot \prod_{h\in N_j\backslash\{i\}} \mathbb{E}\left[e^{rt_{jh}(e_j,e_h)} \middle| e_j \right] \\
\Leftrightarrow & e_i - t_{ij} - \frac{1}{r} \sum_{k\in N_i\backslash\{j\}} \ln \mathbb{E}\left[e^{rt_{ik}(e_i,e_k)} \middle| e_i \right] = e_j + t_{ij} - \frac{1}{r} \sum_{h\in N_j\backslash\{i\}} \ln \mathbb{E}\left[e^{rt_{jh}(e_j,e_h)} \middle| e_j \right] - \frac{1}{r}\ln\alpha_{ij} \\
\Leftrightarrow & t_{ij} = \frac{1}{2}e_i - \frac{1}{2}e_j - \frac{1}{2r} \sum_{k\in N_i\backslash\{j\}} \ln \mathbb{E}\left[e^{rt_{ik}(e_i,e_k)} \middle| e_i \right] + \frac{1}{2r} \sum_{h\in N_j\backslash\{i\}} \ln \mathbb{E}\left[e^{rt_{jh}(e_j,e_h)} \middle| e_j \right] + \frac{1}{2r}\ln\alpha_{ij} \qquad (30) \\
&= \frac{1}{2}e_i - \frac{1}{2}e_j - \frac{1}{2r} \sum_{k\in N_i\backslash\{j\}} \ln T_{ik} + \frac{1}{2r} \sum_{h\in N_j\backslash\{i\}} \ln T_{jh} + \frac{1}{2r}\ln\alpha_{ij} \qquad (31)
\end{aligned}$$

where

$$T_{ik} := \mathbb{E}\left[e^{rt_{ik}(e_i,e_k)} \middle| e_i \right]$$

Then, taking conditional expectations of (30), we have

$$\begin{aligned}
T_{ij} &= e^{r\left(\frac{1}{2}e_i + \frac{1}{2}r\sigma^2 - \frac{1}{2r}\sum_{k\in N_i\backslash\{j\}} \ln T_{ik} + \frac{1}{2r}\ln\alpha_{ij}\right)} \cdot \mathbb{E}\left[e^{r\frac{1}{2r}\sum_{h\in N_j\backslash\{i\}} \ln T_{jh}} \middle| e_i \right] \\
&= e^{r\left(\frac{1}{2}e_i + \frac{1}{2}r\sigma^2 - \frac{1}{2r}\sum_{k\in N_i\backslash\{j\}} \ln T_{ik} + \frac{1}{2r}\ln\alpha_{ij}\right)} \cdot \prod_{h\in N_j\backslash\{i\}} \mathbb{E}\left[T_{jh}^{\frac{1}{2}} \right]
\end{aligned}$$

and

$$\frac{1}{r}\ln T_{ij} = \frac{1}{2}e_i + \frac{1}{2}r\sigma^2 - \frac{1}{2} \sum_{k\in N_i\backslash\{j\}} \frac{1}{r}\ln T_{ik} + \frac{1}{2r}\ln\alpha_{ij} + \sum_{h\in N_j\backslash\{i\}} \ln \mathbb{E}\left[T_{jh}^{\frac{1}{2}} \right].$$

Introducing notation

$$\tilde{T}_{ij} = \frac{1}{r}\ln T_{ij},$$

we have

$$\tilde{T}_{ij} = \frac{1}{2}e_i - \frac{1}{2}\sum_{k\in N_i\setminus\{j\}} \tilde{T}_{ik} + c_{ij}$$

$\Rightarrow$

$$\sum_{j\in N_i} \tilde{T}_{ij} = \frac{d_i}{2}e_i - \frac{1}{2}\cdot(d_i - 1)\sum_{j\in N_i} \tilde{T}_{ik} + \sum_{j\in N_i} c_{ij}$$

$\Rightarrow$

$$\sum_{j\in N_i} \tilde{T}_{ij} = \frac{d_i}{d_i+1}e_i + \frac{2}{d_i+1}\sum_{j\in N_i} c_{ij}$$

$\Rightarrow$

$$\begin{aligned}
\tilde{T}_{ij} &= \frac{1}{2}e_i - \frac{1}{2}\sum_{k\in N_i\setminus\{j\}} \tilde{T}_{ik} + c_{ij} \\
&= \frac{1}{2}e_i - \frac{1}{2}\sum_{k\in N_i} \tilde{T}_{ik} + \frac{1}{2}\tilde{T}_{ij} + c_{ij}
\end{aligned}$$

$\Rightarrow$

$$\frac{1}{2}\tilde{T}_{ij} = \frac{1}{2}\left(e_i - \frac{d_i}{d_i+1}e_i - \frac{2}{d_i+1}\sum_{k\in N_i} c_{ik}\right) + c_{ij}$$

$\Rightarrow$

$$\tilde{T}_{ij} = \frac{1}{d_i+1}e_i - \frac{1}{d_i+1}\sum_{k\in N_i} c_{ik} + c_{ij}$$

Hence, by (30), we have

$$\begin{aligned}
t_{ij} = \frac{1}{2}e_i - \frac{1}{2}e_j - \frac{1}{2}\sum_{k\in N_i\setminus\{j\}}\left(\frac{1}{d_i+1}e_i - \frac{1}{d_i+1}\sum_{k'\in N_i} c_{ik'} + c_{ik}\right) \\
+ \frac{1}{2}\sum_{h\in N_j\setminus\{i\}}\left(\frac{1}{d_j+1}e_j - \frac{1}{d_j+1}\sum_{h'\in N_j} c_{jh'} + c_{jh'}\right) + \frac{1}{2r}\ln\alpha_{ij} \\
= \frac{1}{2}\left(1 - \frac{d_i-1}{d_i+1}\right)e_i - \frac{1}{2}\left(1 - \frac{d_i-1}{d_i+1}\right)e_j + C_{ij} \\
= \frac{1}{d_i+1}e_i - \frac{1}{d_i+1}e_j + C_{ij}.
\end{aligned}$$

$\square$

B.3 Linear Pareto Efficient Transfer Shares for Boundary Correlation Parameters

Proposition 11.

- *For* $\rho = -\frac{1}{n-1}$ *and any network structure* G *such that* $\max_{i \in N} d_i = n - 1$*, let* i^* *be any individual with* $d_{i^*} = n - 1$*. Then a Pareto efficient profile of transfer rules is given by*

$$\alpha_{ji^*} = 1, \quad \alpha_{i^*j} = \alpha_{jk} = 0, \quad \forall j, k \in N \backslash \{i^*\}.$$

- *For* $\rho = 1$ *and any network structure* G*, any profile of transfer rules that satisfies the Kirchhoff Circuit Law as defined below is Pareto efficient:*

$$\sum_{j \in \overline{N}_i} \alpha_{ij} = \sum_{j \in \overline{N}_i} \alpha_{ji} \quad \forall i \in N.$$

Proof. For $\rho = -\frac{1}{n-1}$ and G s.t. $\max_{i \in N} d_i = n - 1$, the profile of transfer rules given above attains zero variance in consumption for each individual, and is thus Pareto efficient. For $\rho = 1$, any profile of transfer rules that satisfies the Kirchhoff Circuit Law achieves the same profile of consumption plan as the null transfer (autarky), which is clearly Pareto efficient. □

B.4. Detailed Specification and Proof for Proposition 8

Specifically, we assume that the correlation between e_i and e_j geometrically decays with social distance between i and j :

$$Corr(e_i, e_j) = \varrho^{dist(i,j)},$$

where the social distance $dist(i, j)$ is formally defined as the length (i.e., the number of links) of the shortest path connecting i and j in network G. For notational simplicity we set $\sigma^2 = 1$.

For tractability, we restrict attention to circle networks with $n = 2m + 1$ individuals. A n-circle consists of n individuals and n links: $G_{i,i+1} = 1$ for $i = 1, ..., n$.[38] For any linked pair $i, i+1$ along a n-circle (with $n \geq 4$), the conditional distribution of e_{i-1} (and similarly for e_{i+2}) is

$$e_{i-1}|_{e_i, e_{i+1}} \sim N\left(\varrho e_i, 1 - \varrho\right).$$

Following a similar argument as in Section 4.2, we obtain the following condition for Pareto efficiency subject to local information constraints:

$$\begin{cases} \alpha_{i,i+1} & = \frac{1}{2}\left(1 - \alpha_{i,i-1} + \varrho \alpha_{i-1,i}\right) \\ \alpha_{i+1,i} & = \frac{1}{2}\left(1 - \alpha_{i+1,i+2} + \varrho \alpha_{i+2,i+1}\right) \end{cases}$$

for all $i \in N$. Then, the unique and symmetric solution for the above system is given by

$$\alpha^*_{ij} \equiv \alpha^{geo}\left(\varrho\right) = \frac{1}{3 - \varrho} \quad \forall G_{ij} = 1.$$

Under α^*, the final consumption for each individual is

$$x_i^{geo}\left(\varrho\right) = \frac{1}{3 - \varrho} e_{i-1} + \frac{1 - \varrho}{3 - \varrho} e_i + \frac{1}{3 - \varrho} e_{i+1}$$

with a variance of

$$Var_{geo,\varrho}\left(x_i^{geo}\left(\varrho\right)\right) = \frac{1 + \varrho}{3 - \varrho}.$$

In comparison, under the symmetric correlation structure in Section 4.2, the condition for Pareto efficiency on a n-circle is

$$\alpha_{i,i+1} = \frac{1}{2}\left[1 - \alpha_{i,i-1} + \frac{\rho}{1 + \rho}\left(\alpha_{i-1,i} - \alpha_{i+2,i+1}\right)\right]$$

with its unique and symmetric solution being

$$\alpha_{ij} \equiv \alpha^{unif}\left(\rho\right) = \frac{1}{3} \quad \forall G_{ij} = 1,$$

which is exactly the local equal sharing rule. This implies a final consumption of

[38] We, for notational simplicity, define individual $n + 1$ to be individual 1, and individual 0 to be individual n.

$$x_i^{unif}(\rho) = \frac{1}{3}e_{i-1} + \frac{1}{3}e_i + \frac{1}{3}e_{i+1}$$

with a variance of

$$Var_{unif,\rho}\left(x_i^{unif}(\rho)\right) = \frac{1+2\rho}{3}.$$

We compare the correlation structures by setting ρ and ϱ to be such that each individual's consumption variance is equalized across the two correlation structures under the global equal sharing rule (which achieves first best risk sharing):

$$x_i^{FB} = \frac{1}{n}\sum_{k\in N} e_k.$$

The consumption variances that this sharing rule implies for the two correlation structures are:

$$\begin{aligned} Var_{unif,\rho}\left(x_i^{FB}\right) &= \frac{1+2m\rho}{2m+1} \\ Var_{geo,\varrho}\left(x_i^{FB}\right) &= \frac{1+2\sum_{k=1}^{m}\varrho^k}{2m+1} = \frac{2\frac{1-\varrho^{m+1}}{1-\varrho}-1}{2m+1}, \end{aligned}$$

The first-best total variances under the two correlations structures are equal if and only if

$$\begin{aligned} Var_{unif,\rho}\left(x_i^{FB}\right) = Var_{geo,\varrho}\left(x_i^{FB}\right) &\Leftrightarrow \frac{1+2m\rho}{2m+1} = \frac{2\frac{1-\varrho^{m+1}}{1-\varrho}-1}{2m+1} \\ &\Leftrightarrow \rho = \rho_m(\varrho) := \frac{\varrho\left(1-\varrho^m\right)}{m\left(1-\varrho\right)}. \end{aligned}$$

Noticing that the total variances without risk sharing at all are both equal to $(2m+1)$ under either correlation structure, setting $\rho = \rho_m(\varrho)$ implies that the total amount of sharable risk is equalized between the two correlation structures. Next we compare the consumption variances given Pareto efficient risk-sharing arrangements subject to local information constraints.

Notice that

$$Var_{unif,\rho}\left(x_i^{unif}(\rho)\right) \leq Var_{geo,\varrho}\left(x_i^{geo}(\varrho)\right) \quad \Leftrightarrow \quad \rho \leq \overline{\rho}(\varrho) := \frac{2\varrho}{3-\varrho}.$$

Hence, whenever

$$m > \frac{(3-\varrho)(1-\varrho^m)}{2(1-\varrho)}$$

we will have $\rho(\varrho) < \overline{\rho}(\varrho)$ and thus $Var^{\rho}_{uni}\left(x_i^{unif}(\rho)\right) < Var^{\varrho}_{geo}(x_i^{geo}(\varrho))$. In other words, fixing ϱ, efficient risk sharing subject to the local information constraint performs strictly better under the uniform correlation setting than under the geometrically decaying setting.

Moreover, the difference can be very stark. As $m \to \infty$,

$$\rho = \rho(\varrho) = \frac{\varrho(1-\varrho^m)}{m(1-\varrho)} \to 0,$$

and thus

$$Var_{unif,\rho}\left(x_i^{unif}(\rho)\right) = \frac{1+2\rho}{3} \to \frac{1}{3}, \quad \text{as } m \to \infty$$

while

$$Var_{geo,\rho}(x_i^{geo}(\varrho)) = \frac{1+\varrho}{3-\varrho} \quad \forall m.$$

When also taking $\varrho \to 1$ (after taking $m \to \infty$), we get

$$\lim_{\varrho\to1}\lim_{m\to\infty} Var_{unif,\rho(\varrho)}\left(x_i^{unif}(\rho(\varrho))\right) = \frac{1}{3},$$
$$\lim_{\varrho\to1}\lim_{m\to\infty} Var_{geo,\rho}(x_i^{geo}(\varrho)) = 1.$$

B.5 Quadratic Utility Functions

With quadratic utility functions $u_i(x_i) = x_i - \frac{1}{2}rx_i^2$, the localized Borch rule requires that

$$\frac{\lambda_j}{\lambda_i} = \frac{E_{ij}\left[u_i'(x_i)\right]}{E_{ij}\left[u_i'(x_j)\right]} = \frac{E_{ij}[1-rx_i]}{E_{ij}[1-rx_j]}$$

$$\Leftrightarrow \quad \lambda_i - \lambda_i r\left(\mu_i + e_i - t_{ij} - \sum_{h\in N_i} E_{ij}[t_{ih}]\right) = \lambda_j - \lambda_j\left(\mu_j + e_j + t_{ij} - \sum_{h\in N_j} E_{ij}[t_{jh}]\right)$$

$$\Leftrightarrow \quad r\left(\lambda_i+\lambda_j\right)t_{ij} = -\left(\lambda_i-\lambda_j\right)+\lambda_i r\left(\mu_i+e_i-\sum_{h\in N_i\backslash j}E_{ij}\left[t_{ih}\left(I_{ih}\right)\right]\right)-\lambda_j r\left(e_j-\sum_{h\in N_j\backslash i}E_{ij}\left[t_{jh}\left(I_{jh}\right)\right]\right)$$

$$\Leftrightarrow \quad t_{ij}=\frac{\lambda_i}{\lambda_i+\lambda_j}\left(\mu_i+e_i-\sum_{h\in N_i\backslash j}E_{ij}\left[t_{ih}\left(I_{ih}\right)\right]\right)-\frac{\lambda_j}{\lambda_i+\lambda_j}\left(\mu_j+e_j-\sum_{h\in N_j\backslash i}E_{ij}\left[t_{jh}\left(I_{jh}\right)\right]\right)-\frac{\lambda_i-}{r\left(\lambda_i+\right.}$$

Postulating a bilateral linear rule:

$$t_{ij}\left(I_{ij}\right)=\alpha_{ij}e_i-\alpha_{ji}e_j+c_{ij}$$

Notice that this is equivalent to specifying $t_{ij}\left(I_{ij}\right)=\alpha_{ij}y_i-\alpha_{ji}y_j+c_{ij}$ as we allow μ_{ij} are simultaneously determined along with:

$$\begin{aligned}
x_i &= \left(1-\sum_{j\in N_i}\alpha_{ij}\right)e_i+\sum_{j\in N_i}\alpha_{ji}e_j+\mu_i-\sum_{j\in N_i}c_{ij} \\
&\equiv \left(1-\sum_{j\in N_i}\alpha_{ij}\right)y_i+\sum_{j\in N_i}\alpha_{ji}y_j+\left(\sum_{j\in N_i}\alpha_{ji}\mu_i-\sum_{j\in N_i}\alpha_{ij}\mu_j\right)-\sum_{j\in N_i}c_{ij}
\end{aligned}$$

Plugging in the postulation,

$$\begin{aligned}
t_{ij} &= \frac{\lambda_i}{\lambda_i+\lambda_j}\left(\left(1-\sum_{h\in N_i\backslash j}\alpha_{ih}\right)e_i+\sum_{h\in N_{ij}}\alpha_{hi}e_h+\sum_{h\in N_i\backslash\overline{N}_j}\alpha_{hi}E_{ij}\left[e_h\right]\right) \\
&\quad -\frac{\lambda_j}{\lambda_i+\lambda_j}\left(\left(1-\sum_{h\in N_j\backslash i}\alpha_{jh}\right)e_j+\sum_{h\in N_{ij}}\alpha_{hj}e_h+\sum_{h\in N_j\backslash\overline{N}_i}\alpha_{hi}E_{ij}\left[e_h\right]\right) \\
&\quad +\frac{\lambda_i\left(\mu_i-\sum_{h\in N_i\backslash j}c_{ih}\right)-\lambda_j\left(\mu_j-\sum_{h\in N_j\backslash i}c_{jh}\right)}{\lambda_i+\lambda_j}-\frac{\lambda_i-\lambda_j}{r\left(\lambda_i+\lambda_j\right)} \\
&= \frac{\lambda_i}{\lambda_i+\lambda_j}\left(\left(1-\sum_{h\in N_i\backslash j}\alpha_{ih}\right)e_i+\sum_{h\in N_{ij}}\alpha_{hi}e_h+\frac{\rho}{1+\left(d_{ij}+1\right)\rho}\sum_{h\in N_i\backslash\overline{N}_j}\alpha_{hi}\cdot\sum_{k\in\overline{N}_{ij}}e_k\right) \\
&\quad -\frac{\lambda_j}{\lambda_i+\lambda_j}\left(\left(1-\sum_{h\in N_j\backslash i}\alpha_{jh}\right)e_j+\sum_{h\in N_{ij}}\alpha_{hj}e_h+\frac{\rho}{1+\left(d_{ij}+1\right)\rho}\sum_{h\in N_j\backslash\overline{N}_i}\alpha_{hi}\cdot\sum_{k\in\overline{N}_{ij}}e_k\right) \\
&\quad +\frac{\lambda_i\left(\mu_i-\sum_{h\in N_i\backslash j}c_{ih}\right)-\lambda_j\left(\mu_j-\sum_{h\in N_j\backslash i}c_{jh}\right)}{\lambda_i+\lambda_j}-\frac{\lambda_i-\lambda_j}{r\left(\lambda_i+\lambda_j\right)}
\end{aligned}$$

In the special case of equal weighting: $\lambda_i = \lambda_j$, we have

$$\alpha_{ij} = \frac{1}{2}\left(1 - \sum_{h\in N_i} \alpha_{ih} + \frac{\rho}{1+(d_{ij}+1)\rho}\left(\sum_{h\in N_i\setminus\overline{N}_j} \alpha_{hi} - \sum_{h\in N_j\setminus\overline{N}_i} \alpha_{hi}\right)\right)$$

$$c_{ij} = \frac{1}{2}(\mu_i - \mu_j) - \frac{1}{2}\left(\sum_{h\in N_i\setminus j} c_{ih} - \sum_{h\in N_j\setminus i} c_{jh}\right)$$

Note that system of linear equations in α is exactly the same one as in Section 4.

B.6 Risk-Sharing Contracts with Ex Post Communication

Consider allowing for a single round of communication after endowment realizations but before transfer payments, where each individual i can send a message $m_{ij} \in \mathscr{M}_{ij}$ to each individual $j \in N\setminus\{i\}$. Then a risk-sharing contract between a linked pair is now a local-state-contingent mechanism that, at each the realized local state, maps (potentially in different ways across each local state) the "locally contractible" (commonly observable) messages to an amount of net transfer. Depending on different specifications of communication protocols, the vector of "locally contractible messages", i.e. messages that can be written into the contract between a given linked pair, may differ. Below is a list of a few simplest communication protocols that are also relevant in reality.

(a) "Global communication": m_{ij} is publicly observable by the whole society. Equivalently, we might as well take $m_{ij} \equiv m_i$ and $\mathscr{M}_{ij} \equiv \mathscr{M}_i$, i.e., each individual can only send a public message that becomes global common knowledge. For example, a global message be thought of as a Tweet, which everyone can observe if he wants to.

(b) "Local announcement": m_{ij} is locally observable by i and i's neighbors. Again, we might as well take $m_{ij} \equiv m_i$ and $\mathscr{M}_{ij} \equiv \mathscr{M}_i$. For example, a local announcement can be thought of as a message i posts on his own Facebook timeline.

(c) "Local comment": m_{ij} is locally observable by j and j's neighbors. For example, a local comment can be thought of as a message i leaves on j's Facebook timeline.

(d) "Private communication": m_{ij} is only privately observable by i and j. A variety of communication technologies such as personal meeting, phone calls, online chats can fit into this category.

For simplicity we equate local common observability with local contractibility in bilateral contracts, and thus the four communication protocols considered above give rise to four different kinds of transfer mechanisms. Let $\mathcal{I}_{ij} \equiv \mathbb{R}^{d_{ij}+2}$ denote the space of local states for linked pair ij and $\mathscr{M} = \times_{i=1}^{n} \mathscr{M}_i$ denote the space of message profile.

(a) With "global communication", the first-best consumption plan $x_i^* (e) = \frac{1}{n} \sum_{k=1}^{n} e_k$ is ex post Nash implementable by a "direct own-endowment mechanism", in which each individual i submits a public report m_i of e_i. Specifically, for the transfer contract $T_{ij} : \mathcal{I}_{ij} \times \mathscr{M} \to \mathbb{R}$ between ij, we may specify it to be contingent on the local state $I_{ij} = (e_k)_{k \in \overline{N}_{ij}}$ and nonlocal reports $(m_k)_{k \in N \backslash \overline{N}_{ij}}$ of $(e_k)_{k \in N \backslash \overline{N}_{ij}}$, which is in particular independent of the i's and j's reports m_i and m_j (of e_i and e_j). Then it is a Nash equilibrium for each individual to report truthfully $m_i (e_i) = e_i$[39], and global information is effectively restored at this Nash equilibrium. In this case, the "informational network" is effectively completed.

(b) With "local announcement", the ex post Nash implementable informational network is effectively given by connecting all 2nd-order neighbors in the physical network G. This can be achieved by a direct mechanism where each individual i submits a report m_i of $I_i \equiv (e_k)_{k \in \overline{N}_i}$. Again, each transfer contract T_{ij} can be specified to be contingent on the local state I_{ij}, as well as i's and j's reports of noncommon neighbors's endowments $(e_k)_{k \in N_i \backslash \overline{N}_j}$ and $(e_k)_{k \in N_j \backslash \overline{N}_i}$. In case of a lie, say, $m_{ik} \neq e_k$ for some $k \in N_i \backslash \overline{N}_j$, this lie is detected by k, and we can specify in the contract T_{ij} a punishment transfer of $|e_k|$ from i to k.[40] It can then be shown that truth-telling is again a Nash equilibrium ex post. Now, suppose a linked pair i and j both "effectively know" e_k for some $k \notin \overline{N}_{ij}$, thanks to truthful local announcements. If $k \in N_i \cup N_j$, then either m_i or m_j, both of which are common knowledge to ij, includes a truthful report of e_k. If $k \notin N_i \cup N_j$, then there must be pentagon subgraph such that k is a diagonal node to link ij in this subgraph. In this case, there is no report on e_k that is commonly known by ij. However, this does not affect risk-sharing efficiency, as the link redundancy in the cycle allows us to specify the contract T_{ij} not to be contingent

[39] This can also be made into a strict Nash equilibrium by incurring punishments of size ϵ for any individual who is found by his neighbors to be lying.

[40] A punishment of size $|e_k|$ is sufficient because, for $\rho \in (0, 1)$, the Pareto efficient transfer shares α are bounded above by 1.

e_k without loss of Pareto efficiency. Lastly, notice that no information about 3rd-order neighbor is possible with local announcements, so no other mechanism can do better.

(c) With "local comment", the ex post Nash implementable informational network is effectively given by connecting all 2nd-order and 3rd-order neighbors. Consider a direct mechanism where each individual i submits to each $j \in N_i$ a report m_{ij} of $I_i \equiv (e_k)_{k \in \overline{N}_i}$.[41] Thus each transfer contract T_{ij} can be specified to be contingent on the local state I_{ij}, all reports received by i and all reports received by j. We may specify the contracts in the following way. If i lies about e_i to j, this is detectable by j, leading to by contract T_{ij} a punishment transfer of amount $|e_i|$ from i to j. This ensures that everyone will truthfully reports his own endowment realization to his neighbors. As a result, each individual i can now observe a truthful report of his 2nd-order neighbors. Now consider a linked pair ij. If i lies to j about e_k for some $k \in N_i \backslash \overline{N}_j$, this is detectable by j because j also observes k's report to i, m_{ki}, which includes a truthful report of e_k. The contract may specify a punishment transfer of amount $|e_k|$ to from i to j, which ensures at a Nash equilibrium the truthfulness of m_{ij} about e_k. Hence, each individual i can now observe a truthful report of his 3rd-order neighbor's endowment e_k, which is included in a report from one of i's 2nd-order neighbor to one of i's (1st-order) neighbor. Now, suppose that both i and j "effectively know" e_k for some $k \notin \overline{N}_{ij}$. If $k \in N_i \cup N_j$, then k submits a truthful report to either i or j, which is commonly observable by i and j, so T_{ij} can be optimally contingent on m_{ki} or m_{kj}. If $k \notin N_i \cup N_j$, there are two possibilities. If k is a 2nd-order neighbor of i (or j with similar arguments), then there must be some $h \in N_i$ that submits a report m_{hi} to i, which is commonly observed by ij and includes truthful report of e_k, so T_{ij} can be optimally contingent on m_{ki}. If k is a 3rd-order neighbor of both i and j, there are three sub-cases. In sub-case 1, ij are both linked to h, of whom k is a 2nd-order neighbor. Then ij commonly observe a report received by h, which includes a truthful report of e_k. In sub-case 2, there exists a path $k \to i$ and a path $k \to j$ that both pass through some $h \in N_k$, but the condition for sub-case 1 does not hold. Then there is no report of e_k that is commonly observed by ij, but h is a diagonal node for link ij in a pentagon subgraph. This does not affect risk-sharing efficiency due to the redundancy of link ij: efficient exposure to e_k can channeled through the two paths

[41] This is different from the direct mechanism considered in (b), as now i can potentially submit different reports m_{ij} to different j.

from h to i and j respectively, and it has been shown in the above arguments that e_k or a truthful report of e_k is commonly observable by the two contracting parties in each transfer contract along the two paths. In sub-case 3, any two paths $k \to i$ and $k \to j$ must be disjoint (except at k), in which case k is the diagonal node to link ij in a heptagon. Again this does not affect risk-sharing efficiency due to the redundancy of link ij: efficient exposures to e_k can still be channeled through the two paths. In particular consider the path $k - i_2 - i_1 - i$, and notice that a truthful report of e_k from i_2 to i_1 is commonly observable by both i_1 and i, thus being contractible. This completes the proof that "local comment" leads to an implementable informational network with effective local observability of 2nd-order and 3rd-order neighbors in the physical network G. Lastly, notice that no information about 4th-order neighbors is possible with local comment, so no other mechanism can do better.

(d) With "private communication", the informational network remains unchanged. This is because when messages are completely private there is no information spillover to any other party. In the meanwhile, the ex post messaging game is a zero-sum game (as the messages are mapped into net transfers). Hence, given each local state I_{ij}, both i and j are guaranteed the value of the ex post game in any Nash equilibrium, so the equilibrium payoffs do not depend on nonlocal endowment realizations. Hence the implementable informational network remains unchanged.

B.7 Detailed Specification and Proof for Proposition 9

In our setting, for a given network G, individual i's Myerson value is defined by

$$MV_i(G) := \sum_{S \subseteq N} \frac{(\#(S)-1)(n-\#(S))}{n!} \cdot \frac{1}{2} r\sigma^2 \left[TVar\left(G|_{(S \setminus \{i\})} \right) + \sigma^2 - TVar\left(G|_S \right) \right]$$

where $\#(S)$ denotes the number of individuals in a subset S of N, and $G|_S$ denotes the subgraph of G restricted to the subset S of individuals. Given the CARA-normal specification, $TVar\left(G|_{(S \setminus \{i\})} \right) + \sigma^2 - TVar\left(G|_S \right)$ is the surplus from risk reduction through i's links in S.

Notice that, given any $S \subseteq N$,

$$TVar\left(G|_{S\setminus\{i\}}\right) - TVar\left(G|_S\right) = 1 - \frac{1}{d_i\left(G|_S\right)+1} + \sum_{k\in N_i\left(G|_S\right)} \frac{1}{d_k\left(G|_S\right)\left[d_k\left(G|_S\right)+1\right]},$$

which is strictly increasing in $d_i\left(G|_S\right)$ but strictly decreasing in $d_k\left(G|_S\right)$ for each $j \in N_k\left(G|_S\right)$. Moreover, for any $k \in N$, $d_k\left(G|_S\right)$ is weakly increasing in S, i.e., $S \subseteq S' \Rightarrow d_k\left(G|_S\right) \leq d_k\left(G|_{S'}\right)$.

Consider any pairwise stable network G under the Myerson-value transfers. Then, if i, j are linked, it must be that

$$MV_i(G) - MV_i(G - ij) \geq c.$$

Fixing ij, for each $S \subseteq N$, we have

$$\begin{aligned} & TVar\left(G - ij|_{S\setminus\{i\}}\right) - TVar\left(G - ij|_S\right) \\ = & \begin{cases} TVar\left(G|_{S\setminus\{i\}}\right) - TVar\left(G|_S\right), & \text{if } j \notin S \\ 1 - \frac{1}{d_i\left(G|_S\right)} + \sum_{k\in N_i\left(G|_S\right)\setminus\{j\}} \frac{1}{d_k\left(G|_S\right)\left[d_k\left(G|_S\right)+1\right]}, & \text{if } j \in S \end{cases} \end{aligned}$$

so

$$\begin{aligned} & \left[TVar\left(G|_{S\setminus\{i\}}\right) - TVar\left(G|_S\right)\right] - \left[TVar\left(G - ij|_{S\setminus\{i\}}\right) - TVar\left(G - ij|_S\right)\right] \\ = & \mathbb{1}\{j \in S\} \cdot \left[\frac{1}{d_i\left(G|_S\right)\left[d_i\left(G|_S\right)+1\right]} + \frac{1}{d_j\left(G|_S\right)\left[d_j\left(G|_S\right)+1\right]}\right] \\ \geq & \mathbb{1}\{j \in S\} \cdot \left[\frac{1}{d_i(G)\left[d_i(G)+1\right]} + \frac{1}{d_j(G)\left[d_j(G)+1\right]}\right] \end{aligned}$$

Averaging over all possible $S \subseteq N$, we get

$$MV_i(G) - MV_i(G - ij) \geq \frac{1}{2} \cdot \left[\frac{1}{d_i(G)\left[d_i(G)+1\right]} + \frac{1}{d_j(G)\left[d_j(G)+1\right]}\right]$$

as

$$\sum_{S\subseteq N} \frac{(\#(S)-1)(n-\#(S))}{n!} \mathbb{1}\{j \in S\} = Pr\{i \text{ arrives later than } j\} = \frac{1}{2}.$$

¿From the perspective of social efficiency, the link ij in G is (strictly) socially efficient if

$$\frac{1}{d_i(G)[d_i(G)+1]}+\frac{1}{d_j(G)[d_j(G)+1]}>2c.$$

Thus we can conclude that, given any pairwise stable network G under the Myerson-value transfers, whenever a link ij is (strictly) socially efficient, it will be present in G, because the increments in both i's and j's private benefits strictly exceed the cost of linking c:

$$MV_i(G)-MV_i(G-ij)>\frac{1}{2}\cdot 2c=c$$
$$MV_j(G)-MV_j(G-ij)>\frac{1}{2}\cdot 2c=c.$$