Network-Constrained Risk Sharing in Village Economies

Pau Milán

This version: January 2016
(June 2015)

Barcelona GSE Working Paper Series

Working Paper nº 912

Network-Constrained Risk Sharing in Village Economies*

Pau Milán†

January 18, 2016

First Draft: June, 2015

Please click here for the latest version of this paper.

Abstract

In this paper I investigate mutual insurance arrangements restricted on a social network. My approach solves for Pareto-optimal sharing rules in a situation where exchanges are limited within a given social network. I provide a formal description of the sharing rule between any pair of linked households as a function of their network position. I test the theory on a unique data set of indigenous villages in the Bolivian Amazon, during the years 2004 to 2009. I find that the observed exchanges across families match the network-based sharing rule, and that the theory can account for the deviation from full insurance observed in the data. I argue that this framework provides a reinterpretation of the standard risk sharing results, predicting household heterogeneity in response to income shocks. I show that this network-based variation in consumption behavior is borne out in the data, and that it can be interpreted economically in terms of consumption volatility. *JEL Classifications: D12, D61, D85, O1, O12*

*I am hugely indebted to Vasco Carvalho, Jan Eeckhout, and Joan de Martí for their continued support, invaluable advice, and excellent guidance throughout this project. I especially want to thank Paul Hooper, Hillard Kaplan, and Michael Gurven for graciously lending the Tsimane' data set from which this paper greatly benefits. I also want to thank Francis Bloch, Christian Brownies, Matt Elliot, Ruben Enikolopov, Sanjeev Goyal, Matt Jackson, Kaivan Munshi, and seminar participants at UPF-Microeconomics Seminar and Cambridge Networks Group for thoughtful comments and discussion. All errors are mine.

†Universitat Pompeu Fabra *e-mail:* Pau.milan@upf.edu

1 Introduction

Vast areas of the developing world rely on informal mechanisms of insurance against random fluctuations in crop-yields and other sources of income. Underdeveloped markets and little financial involvement means that households must often find alternative social arrangements with which to smooth consumption. Typically, these risk sharing arrangements involve the exchanges of goods and services within a village or broader community. A great deal of work has gone into testing the "full risk-sharing hypothesis" under which, if communities are indeed hedging risk efficiently, idiosyncratic and independent movements in income should not correlate with fluctuations in consumption.[1] While this test is widely accepted now as the standard approach to test full insurance, it can only provide evidence for or against Pareto-optimal allocations; it fails, however, to provide an accurate alternative characterization below efficiency. Moreover, most empirical work on the subject has repeatedly rejected full risk sharing in a number of different contexts ranging from India to Tanzania and including Thailand, Peru, and many others.

In this paper I present a complementary interpretation of the risk-sharing test that provides a more detailed account of the type of behavior we might observe when we reject full risk sharing. In particular, I account for local network interactions by constraining Pareto-optimal allocations to a situation where exchanges are limited within a given social arrangement. As a theoretical contribution, I provide a formal description of the sharing rule between any pair of linked households strictly as a function of their network position. I structurally estimate the sharing rule against a unique data set of indigenous communities in the Bolivian Amazon, and I show that this description of sharing behavior does a remarkably good job at describing observed transfers across families. I argue that this framework provides a reinterpretation of the standard risk sharing results, predicting household heterogeneity in response to income shocks. I show that this network-based variation in consumption behavior is borne out in the data, and that it can be interpreted economically in terms of consumption volatitliy. Finally, I show the theory can account for the level of risk sharing observed in the data.

The current framework provides a general approach to modeling mutual insurance organized around local risk sharing groups. It generalizes recent work that has approached within-group insurance largely as an empirical question.[2] Rather than taking groups as separate, perfectly insured communities, I allow for a fully general network with interconnected sharing groups that are specific to each household. I argue that, in this environment, not defining the relevant local sharing group biases the results of classical tests of full insurance. More importantly, I show that controlling for this bias does not

[1] Cochrane (1991) best explained this approach as the cross sectional equivalent to Hall's permanent income hypothesis test, which regressed consumption growth rates over time on ex-ante variables (Hall, 1978). Indeed, under complete borrowing and lending opportunities consumption should not respond, over time, to forecastle shocks, just as it should not respond to idiosyncratic shocks across households under full insurance.

[2] Hayashi et al. (1996) consider whether extended families can be viewed as collective units sharing risk efficiently. Mazzocco and Saini (2012), for instance, argue that the relevant sharing group in India is the caste, rather than the village. Munshi and Rosenzweig (2009) also find that the caste is the relevant group to explain migration patterns in rural india. More recently Attanasio et al. (2015) test for efficient insurance within extended families in the U.S.

eliminate the correlation between consumption and income across households: the network structure generates underlying heterogeneity in sharing behavior, which implies that households' income affects consumption even after appropriately controlling for local aggregates.

I solve a constrained welfare problem in which transfers are limited along a given social structure. The restriction on exchanges means that whatever a household receives from its neighbors cannot be shared further down the network; that is, I assume that income can be split and shared only amongst immediate neighbors. This assumption is meant to capture the relatively low levels of intermediation, relative to direct exchanges, that occur in these types of subsistence economies, where mostly crops and other perishable goods are traded.[3] Alternatively, even when risk-sharing involves the transfer of cash as well, urgent liquidity needs often means households cannot immediately access distant funds that must first be intermediated by the network.[4] While I take the lack of any intermediation as a simplifying assumption, I also show that most results can be sustained when allowing for greater movement of funds.[5]

In this context of no intermediation I solve for the non-contingent (or fixed) sharing rules that maximize welfare. These type of sharing rules specify a fraction of each household's income consumed by each sharing partner, where this fraction is constant across all states of the world. In Section 2 I discuss the implications of these type of sharing rules in the context of a very simple example, and I show that they allow me to isolate network effects from income distributions in order to obtain simple predictions on sharing behavior across households.

Of course, various other explanations have been provided to account for the failure of full risk sharing in village economies. For instance, a number of papers have argued that incomplete information, limited commitment, heterogeneous preferences, or the presence of outside markets are all capable of generating inefficiencies in mutual insurance mechanisms.[6] In this context, it is worth asking why it makes sense to model social networks as a constraint on the classical welfare problem. Evidence suggests that individuals select into particular social arrangements precisely to mitigate informational frictions and to guarantee compliance, so that within these social arrangements mutual insurance mechanisms function rather well.[7] Moreover, there is strong evidence that informational frictions

[3] Hooper (2011) for instance mentions it is quite rare to observe the same good exchange hands twice within Tsimane' communities. Similarly, Chiappori et al. (2013), Kinnan and Townsend (2012) and Udry (1994) document that an overwhelming share of the economies they study in Thailand and Nigeria are formed by crops, livestock and other perishable goods.

[4] You can think of this assumption as the complementary version of the assumptions driving the model of Ambrus et al. (2014). In that model, funds can travel indefinitely along the network, but each edge has some exogenous capacity constraint that limits the amount of funds it can intermediate. In this case, intermediation is ruled out, but the amount of funds along any given edge is endogenized. These type of limited interactions are also studied by Bourlés et al. (2015) in the context of altruism in networks, with very different implications.

[5] In Section 8.2 I show how to extend results to a general case with network intermediation. Notice that without some limit on how far funds can exchange hands along the network, the welfare problem is unconstrained and full insurance obtains as the unique outcome.

[6] See for instance Incomplete information (Udry, 1994); Limited Commitment (Ligon et al. 2002); Heterogeneous preferences (Schulhofer-Wohl 2015, Mazzocco & Saini 2015); Outside Markets (Munshi & Rosenzweig 2014, Galeotti et al. 2015, Saidi 2015).

[7] Munshi (2014) describes the general tendency of households to arrange into particular social patterns that avoid

within these social spaces are relatively unimportant.[8] In this paper I abstain from considering the forces that shape particular social networks. Instead, I take them as given and study the type of efficient outcomes we expect within these restricted environments.

My first main result relates constrained-efficient transfers to a global measure of households' relative importance in the network. This measure reflects a household's direct and indirect interactions along the entire network. In this manner, the proposed measure bears some similarity to previous statistics — for instance Katz-Bonacich or PageRank — that capture higher-order dependencies as they feed back along a given network of connections. The particular flavor of this network measure has to do with the tradeoff faced by the planner between variance and covariance considerations. I find that, for any network, the constrained-efficient exchanges across any pair of households follows a simple relationship between the sender and receiver's network measures.

I show that this framework redefines the relationship between consumption and income in a setting of partial insurance on a network. This has important implications for the standard risk sharing tests. Specifically, the model provides heterogeneous predictions on household's response to income shocks that emerge from households' different positions in the network. This provides insight into the varied insurance possibilities of households when full risk sharing is rejected. I show that, in certain scenarios, these values can be mapped to important economic features, such as consumption volatility.

Having established the theoretical results, I examine a unique data set of about 250 households in 8 different indigenous villages in the Bolivian Amazon basin, during the years 2004 to 2009. This data is particularly well suited for my analysis as it provides information on the caloric exchanges across pairs of households over time.[9] This allows me to structurally estimate the model by fitting the theoretical relationship between network position and exchanges at the edge level. Moreover, compared to other models of risk sharing networks, I can estimate the model at a much finer level of variation (i.e. using edge-level data), and separately from aggregate considerations on consumption growth. I find that the empirical flows across connected dyads indeed respond to the network structure as the model prescribes. I show that once we account for this restriction on the planner's problem, we can effectively explain all the variation in consumption that correlates with households' income. Finally, I also test the model's implication on network-based heterogeneity of households' response to own income shocks, and I find that the data exhibits the same type of variation that the model prescribes. I do this by constructing a couple of tests that can be applied to many other data sets that include network and income data; it can be tested on a wide range of empirical settings. The results suggest that previous failures of full-risk-sharing tests are best understood by invoking restrictions on bilateral exchanges.

certain commitment issues. On the theoretical side, a long list of papers have studied the type of networks that might emerge under limited commitment and similar frictions. Bramoullé & Kranton (2007), Jackson et al. (2012), and Ambrus et al. (2015) are just a few.

[8]In his work on Nigerian communities, for instance, Udry (1994) argues that loan arrangements are very informal, with no collateral, explicit interest rates or repayment dates, and that households know each other well. Hooper (2011) also finds similar evidence of strong informational flows in the Tsimane' networks of Bolivia that I study in this paper.

[9]Exchanges are measured in calories: food is the primary source of income and trade for subsistence economies like the Tsimane' communities studied in this paper.

Related Literature

The distribution of uncertainty along social ties has, in the past several years, drawn a lot of interest from economists. Starting with Bramoullé and Kranton (2007a,b) and Bloch et al. (2008), a number of recent contributions — such as Jackson et al. (2012), Billand et al. (2012), Ali and Miller (2013a,b), and Ambrus et al. (2014) — have focused on enforcement concerns and the role of social capital in sustaining cooperative behavior. Most of these studies assume networks serve a dual role as both medium of exchange and social collateral, delivering efficient and stable structures for a set of exogenous, and fixed, bilateral transfers. In other words, most of these papers assume a sharing behavior and find networks that sustain it. Bramoullé and Kranton's (2007a,b) model, for instance, assumes that a connected component equally distributes its surplus independent of the social structure, so that inequality is ruled out. Billand et al. (2012) also assume a sharing behavior whereby high-income households transfer a fixed amount to low-income neighbors. I take a different view that abstracts from enforceability considerations altogether and instead provides an endogenous prediction of efficient transfers along a network. The focus on the distribution of surplus, and away from enforceability, appears most recently in work by Ambrus et al. (2015) that studies cross-group incentives for social investments. However, their concern has to do with network formation, so they also assume some exogenous split of surplus: bilateral exchanges are assumed to split the total surplus according to the Shapley value, which, in the particular setting they focus on, reduces to equal sharing. Perhaps closest in spirit is the work by Ambrus et al. (2014) that similarly refrains from assuming, a priori, the sharing pattern across connected pairs; they otherwise assume a distribution of "link values" that are perfectly substitutable with consumption, so that coalition-proof transfers are, again, ultimately determined from outside the model. The current paper refrains from engaging with these difficult strategic considerations, and instead solves for a simple constrained-efficient, network-based sharing rule that provides a number of testable implications.

On the empirical side, this paper joins the ranks of a long strand of research devoted to the estimation, and interpretation, of risk-sharing patterns in data. While newer data sets have begun to include social surveys that allow us to test network models directly, the empirical risk-sharing literature has a longer tradition, and one that, with occasional exceptions, has overwhelmingly insisted that communities operate below efficiency. The work of Mace (1991), Cochrane (1991), and Townsend (1994) provided the theoretical foundations for measuring correlations between household income and consumption, which, by now, has become the hallmark of all empirical tests on risk sharing. Since then, a healthy number of studies have sprung up to investigate one or another economic dimension of risk-sharing communities — from the impact of kinship ties on credit constraints in the Philippines (Kinnan and Townsend, 2012) to the decreased social mobility induced by local sharing along caste lines (Munshi and Rosenzweig, 2009). Whatever the particularities, all these studies perform the standard test of full risk-sharing and, together, deliver a cogent narrative that by and large strays away from efficiency. For instance, Ligon (1998) studies a private information alternative to the complete market model and rejects full insurance in rural south India. Fafchamps and Lund (2003) famously reject full insurance for Philippine communities and show that the extent of risk sharing is limited by the extent of interpersonal networks. Mazzocco and Saini (2012) reject full insurance for indian data at the village but not at the caste level, while Munshi and Rosenzweig (2009) reject efficiency at the caste level as well. On their study of investment decisions under exogenous income shocks to networks in rural Mexico, Angelucci et al. (2015) however find that they cannot reject full insurance within extended families. The list is long, though, and more often than not signals of full risk-sharing are absent from a wide range of settings.[10] More importantly, Saidi (2015) studies credit demand in the same Tsimane' indigenous communities that I study and also rejects full insurance.

Finally, the paper also relates to a number of studies that have sought to provide a direct explanation for the repeated failure of efficiency in data. For instance, Ligon et al. (2002) model optimal contracts under limited commitment. They estimate their model on three separate indian villages, and argue that this type of transaction cost accounts for the magnitude of departure from full insurance. More recently, a couple of studies have argued that heterogeneous risk preferences might force an interpretation of full risk-sharing test that is far too pessimistic. (Mazzocco and Saini (2012) or Schulhofer-Wohl (2011)). Schulhofer- Wohl (2012), for instance, argues that if households' variation in risk preferences are cyclical then not accounting for these explicitly introduces an omitted variable bias that pushes the coefficient of own income upwards, leading to false rejections of full insurance. Mazzocco and Saini (2012) have similarly developed empirical tests for heterogeneous preferences and provided a modified empirical procedure to test for efficiency. Most importantly, Fafchamps and Lund (2003) address the failure of efficient insurance in the data by invoking the role of gifts and remittances in risk-sharing and reject mutual insurance at the village level, suggesting instead that households receive transfers from a network of family and friends. Although they don't model network flows explicitly, their findings serve as the principal motivation for this paper.

The remainder of the paper is organized as follows. Section 2 goes over the standard test of full insurance and argues how the current setup affects this estimation procedure using a simple example. Section 3 introduces the theoretical framework, solves for the efficient sharing behavior, and provides implications for the test of full insurance. Section 4 provides background information about the data and summary statistics. In Section 5 I structurally estimate the model, and draw a number of testable implications for risk-sharing tests. Section 6 concludes.

2 Network Constrained Risk Sharing: A Simple Example

In this section I present the canonical model of full risk sharing and I describe the empirical approach that emerges from it to test full insurance from data. I then describe the main assumptions behind

[10] As Schulhofer-Wohl (2012) reminds us, full insurance has been rejected in data from the United States (Attanasio and Davis 1996; Cochrane 1991; Dynarski and Gruber 1997; Hayashi et al. 1996), Côte d'Ivoire (Deaton 1997), India (Munshi and Rosenzweig 2009; Townsend 1994), Nigeria (Udry 1994), and Thailand (Townsend 1995). Mace (1991) does not reject efficiency in U.S. data, but Nelson (1994) overturns this result.

this paper and how it refines the concept of sharing groups. I use a very simple example to describe the sharing rules I obtain, and I explain what they claim about the distribution of insurance across the population. Finally, I present the implications of this model on the standard risk-sharing tests and I show that, 1) not defining the appropriate local sharing groups generates biased estimators, and 2) network asymmetries generate varying predictions on the impact of income shocks on consumption, which in turn provides a network story behind the rejection of full insurance. Section 3 then generalizes all these arguments to a full fledged model with an arbitrary network and a general income process for households.

2.1 Canonical Model

The classical risk sharing models of Cochrane (1991), Mace (1991), and Townsend (1994) solve for the ex-post pareto-optimal allocations by defining a planner problem as follows,

$$\max_{c_i(\omega)} \sum_{\omega} \pi\left(\omega\right) \sum_{i} \eta_i \, u_i\left(c_i\left(\omega\right)\right)$$

where $\pi\left(\omega\right)$ represents the probability of state ω and where η_i represents $i's$ Pareto weight in the welfare function. This problem is subject to the constraint that total consumption not exceed total income in any state of the world, or that $\sum_i c_i\left(\omega\right) \leq \sum_i y_i\left(\omega\right)$ for all $\omega \in \Omega$, .[11] The first order conditions yield the well known full insurance equations known as Borch's rule. It states that the ratio of marginal utilities across any two agents is constant across states. Formally, we can solve for the problem above and, for any two households i and j, obtain the following expression,

$$\frac{u_i'\left(c_i\left(\omega\right)\right)}{uj'\left(c_j\left(\omega\right)\right)} = \frac{\eta_j}{\eta_i}\,, \;\; for\; all\; \omega \in \Omega \tag{1}$$

This expression has been used to develop a popular test of full insurance. Indeed, equation (1) states that, under full risk sharing, consumption should not respond to idiosyncratic shocks after controlling for aggregate shocks. The following type of regressions,

$$log(c_{it}) = \alpha_i + \beta_1 log(y_{it}) + \beta_2 log(\bar{y}_t) + \epsilon_{it} \tag{2}$$

where $\bar{y}_t$ represents aggregate income, have been used to test for efficient outcomes, in which case $\beta_1 = 0$ and $\beta_2 = 1$. Time and again, β_1 is found to be positive and significant and β_2 below one. Unfortunately, not much can be learned from these results other than the existence, or not, of full insurance. The following approach attempts to give a more nuanced understanding of the type of sharing behavior that might be generating these estimates.

[11]Nondecreasing utility functions on consumption means that the constraint will hold with equality.

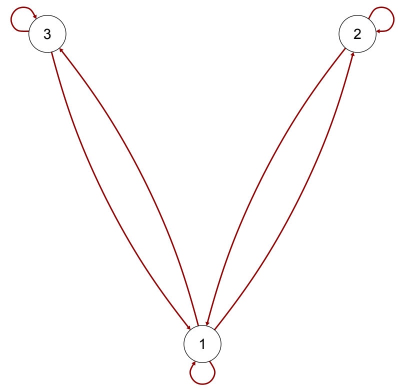

Figure 1: A Simple Risk Sharing Economy

2.2 Overlapping Sharing Groups

An important feature of the classical risk sharing model above is that all households form part of the same risk sharing group. In this paper, I relax this assumption by considering the possibility that mutual insurance is local. This allows me to capture a number of relevant intermediation costs that might make it impossible to define a unique sharing group.[12] If these motives are strong, households can only access local risk sharing groups defined by their immediate neighbors (or trading partners).[13]

To fix ideas, consider the economy presented in Figure 1 where households 2 and 3 can only trade with household 1. For simplicity, imagine all households obtain a random income realization $y_i(\omega)$ that is i.i.d. from some distribution $F\left(\mu,\sigma^2\right)$. The lack of intermediation means households must access different, overlapping risk sharing groups — for instance, the risk sharing group of household 2 consists of households 2 and 1 only. Let α_{ij} represent the fraction of $j's$ income consumed by i. The situation of this economy can be written as follows,

$$\begin{array}{lllllll} c_1(\omega) & = & \alpha_{11}y_1(\omega) & + & \alpha_{12}y_2(\omega) & + & \alpha_{13}y_3(\omega) \\ c_2(\omega) & = & \alpha_{21}y_1(\omega) & + & \alpha_{22}y_2(\omega) & & \\ c_3(\omega) & = & \alpha_{31}y_1(\omega) & + & & + & \alpha_{33}y_3(\omega) \end{array} \tag{3}$$

This formulation provides a very tractable way to define the risk sharing rule in this economy by expressing consumption explicitly in terms of bilateral transfers, α_{ij}. Notice that we can describe the canonical model above as the particular case where households 2 and 3 are able to trade with each other because they are directly connected.[14] In this case, all households clearly access the same risk sharing group and efficiency obtains.

[12]For instance, the Tsimane' communities that I study in this paper transfer highly perishable goods - mostly prepared food and game. In other contexts where risk sharing involves the transfer of cash as well, urgent liquidity needs means individuals cannot immediately access distant funds that must be intermediated.

[13]In Section 8.2 I show how to extend the model to all levels of intermediation. Notice that if intermediation is sufficiently high, all households access the same risk sharing group and efficiency obtains as above.

[14]Or, alternatively, by intermediating through household 1.

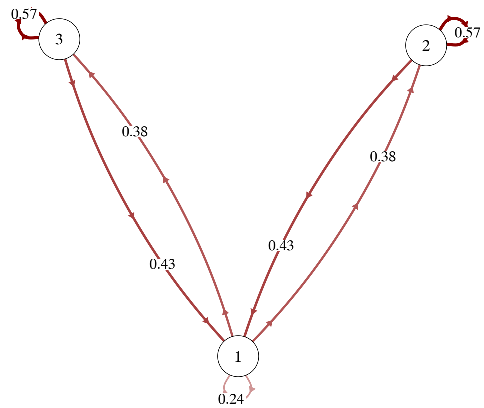

Figure 2: The Sharing Rule of a Simple Economy

In this paper I solve generically for the set of non-contingent sharing rules that maximize welfare in this setup with no intermediation. A non-contingent sharing rule means that the fraction α_{ij} of $j's$ income consumed by i is constant across all states ω.[15] In this paper I show how to solve analytically for this type of sharing rule for any given network. As an example, consider the economy of figure 1, and, in order to make the argument as simple as possible, set all Pareto weights η_i equal and set $\mu^2 = \sigma^2$. Applying the main theoretical result of this paper we obtain the following simple description between a household's consumption and the incomes of its relevant sharing group:

$$\begin{array}{lllllll} c_1(\omega) & = & \frac{5}{21} y_1(\omega) & + & \frac{9}{21} y_2(\omega) & + & \frac{9}{21} y_3(\omega) \\ c_2(\omega) & = & \frac{8}{21} y_1(\omega) & + & \frac{12}{21} y_2(\omega) & & \\ c_3(\omega) & = & \frac{8}{21} y_1(\omega) & + & & + & \frac{12}{21} y_3(\omega) \end{array} \tag{4}$$

This situation is depicted in Figure 2. A great deal can be gleaned already from this very simple example. Notice that households 2 and 3 share a larger fraction of their income with 1 than 1 shares with them ($\frac{9}{21} > \frac{8}{21}$); still household $1's$ relevant sharing group is larger and as a result 1 consumes much less of its own income than 2 or 3 ($\frac{5}{21} < \frac{12}{21}$). Moreover, it is easy to show that consumption volatility associates positively with this value, so that household 1 (with a lower coefficient) obtains a less volatile consumption stream than households 2 and 3. This setup therefore provides network-based heterogeneity on households' response to own income shocks and relates it to the distribution of risk sharing opportunities.

[15]In 8.1 I discuss the alternative assumption that sharing rules are contingent — i.e. $\alpha_{ij}(\omega)$. I show how it restricts the set of states for which the efficient condition (1) holds, and I demonstrate the inherent difficulty in isolating general network effects from particular income realizations for these type of contingent sharing rules. I also provide some empirical evidence that suggests informal exchanges in village economies might be closer to a fixed (or non-contingent rule) than to an extremely flexible sharing rule.

2.3 Risk Sharing Regressions under Local Insurance

Ultimately, these predictions generate enough information on household consumption to provide reasonable explanations for the rejection of full insurance. I consider how this affects the empirical tests of risk sharing described in the previous section. Let us stick to the simple economy in Figure 1 and consider rewriting equations (3) in the form of the classical risk-sharing specification of equation (2) with a common aggregate income term,

$$\begin{aligned}
c_{1t} &= (\alpha_{11} - \alpha_{12})\, y_{1t} + \alpha_{12}\bar{y}_t + \epsilon_{1t} \\
c_{2t} &= (\alpha_{22} - \alpha_{21})\, y_{2t} + \alpha_{21}\bar{y}_t + (\epsilon_{2t} - \alpha_{21} y_{3t}) \\
c_{3t} &= (\alpha_{33} - \alpha_{31})\, y_{3t} + \alpha_{31}\bar{y}_t + (\epsilon_{3t} - \alpha_{31} y_{2t})
\end{aligned}$$

These equations reflect three important themes of this paper: 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' position in the network, 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.[16] The classical risk sharing test in (2), pools these equations and obtains a unique estimate for β_1; given the previous discussion we expect this estimate to be different from zero and positive. In the first column of Table 1 I show the estimates for the simple example of Figure 1 for simulated data.[17] As expected, β_1 is statistically significant and close to 0.2, while the coefficient on the common aggregate income term, β_2 , is statistically lower than 1.

In order to isolate the network effect from the bias in β_1, consider estimating (2) with the relevant local sharing group instead. In this case, estimates are no longer biased, 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. To see this rewrite again equations (3) in the form of (2), but now we allow for household-specific aggregates, $\bar{y}_{it}$, that sum over the incomes of $i's$ local sharing group. In this case we have

$$\begin{aligned}
c_{1t} &= (\alpha_{11} - \alpha_{12})\, y_{1t} + \alpha_{12}\bar{y}_{1t} + \epsilon_{1t} \\
c_{2t} &= (\alpha_{22} - \alpha_{21})\, y_{2t} + \alpha_{21}\bar{y}_{2t} + \epsilon_{2t} \\
c_{3t} &= (\alpha_{33} - \alpha_{31})\, y_{3t} + \alpha_{31}\bar{y}_{3t} + \epsilon_{3t}
\end{aligned}$$

Because aggregates are now household-specific, the additional terms in the error term disappear and we obtain unbiased estimators. Notice, however, that coefficients to own income are different from zero so long as $\alpha_{ii} - \alpha_{ij} \neq 0$. This implies that the pooled regression will again deliver positive coefficient, β_1, even with the appropriate local aggregates. I present the results to this local sharing group version

[16] If incomes are positively correlated, then imposing a common aggregate variable biases estimates upwards. Schulhofer-Wolf similarly uncovers a bias in the classical risk-sharing specification that comes from heterogeneity in income preferences. Here, the heterogeneity is induced by positions in social structures. In any case, as I show below, we can adjust for the bias and still expect positive coefficients to income in this setup.

[17] I simulate log-normal income data for all three households with $t = 100,000$ and I obtain household consumption as indicated by the sharing rule in (4). I then run the standard risk-sharing regression on logged data, controlling for household fixed effects.

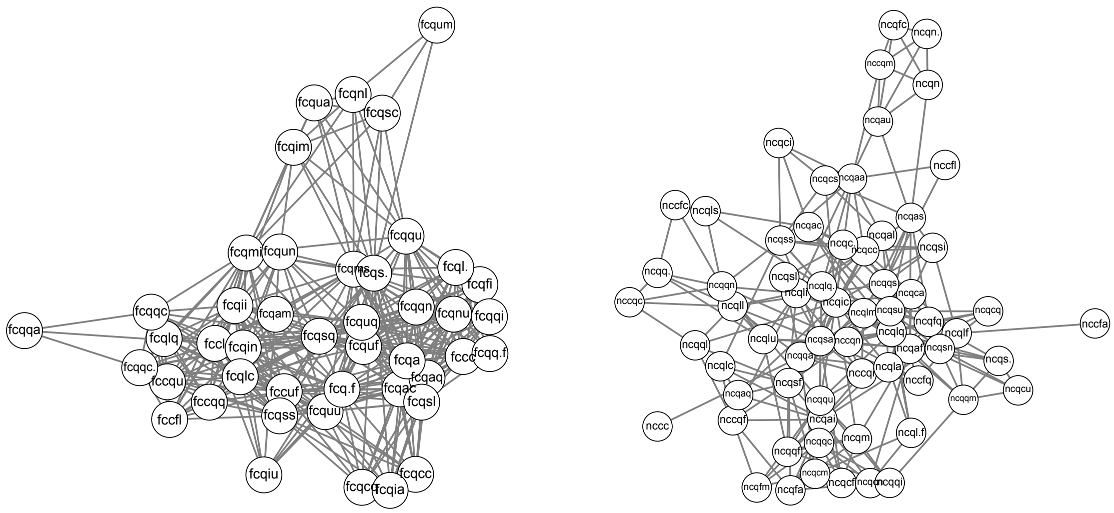

Figure 3: Two Tsimane' Villages: (a) Kinship Network (b) Trade Network

of equation (2) in the right column of Table 1. Again, the coefficient to income is positive, as expected, although estimates decrease by one order of magnitude.

Real world social structures are usually far more complicated than these simple examples. Figure 4 plots two of the networks I build from data in one of eight indigenous communities I study in this paper; the one on the left is built from kinship data and the one on the right on trade data.[18] These networks are orders of magnitude more complicated. Still, I show that the arguments above can be extended generically for any network and general income process across households. Moreover, this unique data set contains information on the transfer of food across households over time, so I am able to structurally estimate the endogenous, network-based sharing rule that I derive in this paper. I find it does a remarkably good job at describing the patterns of exchange across households in these subsistence economies.

3 The Model

I study an economy in which households face uncertainty about their income realizations, but may redistribute incomes through a given network of social connections. I characterize efficient transfers as a function of households' position in the network when the movement of funds is restricted. In section 3.1 I describe the theoretical setup. In section 3.2 I solve for the constrained-efficient set of transfers and describe how they relate to the underlying network. In section 3.3 I provide certain properties of the sharing rule and describe its behavior more closely for some simple structures.

[18]Refer to section X for a detailed description of the types of networks constructed and a discussion on the relative merits of each. Refer to Figures 1 and 2 in the appendix for a visual plot of all villages for each of the network types.

3.1 Setup

Consider a population of size N arranged in a network $L = (V, E)$, consisting of a set V of households (vertices) and a set E of pairs of elements of V that represent links (edges) across these households. I assume the network is undirected, so that the pair of vertices in E is unordered. It is also useful to define an alternative characterization of this social structure by an adjacency matrix G, where $g_{ij} = 1$ if and only if $\{i, j\}, \{j, i\} \in E$. Each connection can represent a friendship, kinship relation, or other type of social connection between the two parties involved. We will refer to i's neighborhood as the subset of N defined by $N_i = \{j \in N \mid e_{ij} \in E\}$. The degree of a vertex i measures the number of connections of i and is defined as the cardinality of N_i.

All households face risky endowments. Denote the vector of random endowments by $\mathbf{y} = (y_i)_{i \in N}$, drawn from some joint distribution F with mean μ and variance σ^2. I assume a common covariance between the incomes of any two agents and denote it by $\rho \neq 0$. I assume throughout that $|\rho| < \sigma^2$ so that incomes are not perfectly correlated.

Households share incomes along a social network, so that consumption levels will differ, in general, from their income realizations. Incomes can only be exchanged once, so that households consume incomes from immediate neighbors.[19] The shares of neighboring endowments consumed by a given household are defined ex-ante and are non-contingent. Together this implies that a household's consumption in state ω can be defined as a linear combination of neighbors' incomes as,

$$c_i(\omega) = \sum_j g_{ij} \alpha_{ij} y_j(\omega) \tag{5}$$

where α_{ij} represents the share of $j's$ endowment that is consumed by i. We will also define $\alpha_i = (\alpha_{ij})_{j \in N_i}$ as the vector of $i's$ incoming shares. By defining the "sharing matrix" $\mathbf{A}$ as $\mathbf{A}_{ij} = g_{ij}\alpha_{ij}$, we can express equation (5) in matrix form in the following way, $\mathbf{c} = \mathbf{A}\mathbf{y}$, where I drop the explicit dependency on ω from now on for notational convenience. Of course, the elements of $\mathbf{A}$ represent percentage claims on neighboring incomes and must therefore satisfy a feasibility condition that all claims on a given endowment sum to 1, which can be expressed as $\mathbf{1} = \mathbf{A}'\mathbf{1}$. Finally, I assume all households have quadratic utility functions:

$$u(c_i) = c_i - \frac{1}{2}\gamma c_i^2$$

where γ is the common coefficient of risk aversion.

I now define the planner problem and provide a short discussion on the particular form of the objective function and the constraints.

Definition 1. *The planner problem is defined as,*

[19] In Section 8.2 I show how this assumption can be relaxed to allow for network intermediation.

$$\underset{\{\alpha_{ij}\}_{ij}}{max}\ \mathbb{E}\sum_i u\left(c_i\right) = \underset{\{\alpha_{ij}\}_{ij}}{min}\ \sum_i \left(\mu^2\left(\sum_j g_{ij}\alpha_{ij}\right)^2 + \sigma^2\sum_j g_{ij}\alpha_{ij}^2 + \rho\sum_{k\neq j} g_{ik}g_{ij}\alpha_{ij}\alpha_{ik}\right) \tag{6}$$

subject to $\alpha_{ij} \geq 0$ *for all* $i, j \in N$ *and that* $\sum_i \alpha_{ij} = 1$ *for all* $j \in N$

The form of equation (6) exploits the linear mean-variance tradeoff of expected utility: the first term in brackets corresponds to the squared mean of consumption, while the next two terms correspond to the variance of consumption.[20] The constraints on the planner problem reflect the standard feasibility conditions that shares are positive and sum to one. Finally, since the sum is convex in shares and the constraint set is linear, the maximization is a convex program and the first order conditions completely characterize the optimal solution. In the next section I define these optimality conditions and explore the type of network interactions that are contained in them. I then provide the general solution for any network $\mathbf{G}$ and a general class of distributions F.

3.2 Constrained-Efficient Network Flows

Having defined the economy and the welfare problem in the previous section, we are now ready to obtain a description of the sharing rule for any given network. To do this in a way that clarifies the type of network interactions that emerge, I first analyze the planner's optimality condition in some detail. The first order conditions of (6) defines the share α_{ij} that i receives from j (for each pair $i, j \in N$) as,

$$\alpha_{ij}^\star = g_{ij}(\Lambda_j - \Psi\sum_k g_{ik}\alpha_{ik}^\star) \quad for\ all\ i, j \in N \tag{7}$$

where $\Lambda_j = \frac{\lambda_j}{2(\sigma^2-\rho)}$, and $\lambda_j > 0$ is the multiplier for $j's$ constraint, and where $\Psi = \frac{\mu^2+\rho}{\sigma^2-\rho} > 0$ captures the shape of the income distribution. It is worthwhile to examine equation (7) in some detail. First of all, notice that if i and j are not connected, $g_{ij} = 0$ and i consumes none of $j's$ income. Instead, if $g_{ij} = 1$ then the fraction of $j's$ income consumed by i depends on two terms. The first term, Λ_j, captures the relationship among all of $j's$ shares, as governed by $j's$ constraint, $\sum_i \alpha_{ij} = 1$. It enters positively because a drop in one of $j's$ shares (holding everything else constant) would increase Λ_j, and thus force all of $j's$ shares up to meet the constraint. As such, this term effectively connects all of the first order conditions pertaining to j. For instance, if no other effect existed, Λ_j would set all of $j's$ shares equal to each other. However, in most situations $j's$ shares are not equal, given the second term in (7). This second term determines how all shares received by i affect α_{ij} — the more i receives from some neighbor k the less it receives from j (and vice versa), where the constant Ψ mediates the

[20] Notice we can write the planner problem as, $\mathbb{E}u\left(c_i\right) = \mathbb{E}\sum_i\left(c_i\right) - \frac{1}{2}\gamma\sum_i\left(\mathbb{E}\left(c_i\right)^2 + var\left(c_i\right)\right)$ and the first term drops out because aggregate consumption must equal aggregate income by the constraints— i.e. $\sum_i c_i = \sum_i y_i$. Therefore, the planner problem reduces to minimizing $\sum_i\left(\mathbb{E}\left(c_i\right)^2 + var\left(c_i\right)\right)$ which corresponds to the expression in Definition 1. In the appendix I show this problem corresponds to the minimization of expected inequality and I relate it to other similar results for CARA utility in Ambrus et al. (2015).

strength of this response. The value of Ψ captures the relative variance and covariance considerations of the planner: as covariance effects increase (and Ψ increases), $j's$ shares respond more to the value of other shares.[21] To sum up, the share of $j's$ income consumed by i responds, on the one hand, to all shares coming from j (through Λ_j) and, on the other hand, to all shares going to i (through Ψ).

More generally, the second term in (7) defines a recursive relationship for α_{ij}. Cutting through the recursivity allows us to reframe the optimality condition (7) in terms of the constraints Λ_k as follows[22],

$$\alpha_{ij}^{\star} = g_{ij}\left(\Lambda_j - \frac{\Psi}{1+\Psi d_i}\sum_k g_{ik}\Lambda_k\right) \quad for\ all\ i,j \in N \tag{8}$$

Given the arguments above, Λ_k connects all of household $k's$ optimality conditions via $k's$ constraint (if α_{ik} decreases, then α_{jk} increases for all j connected to k, holding everything else constant). Therefore, equation (8) expresses α_{ij} not as a function of all shares that i receives (as in (7)), but instead as a function of the full set of interactions for each of $i's$ partners. That is, it contains all the indirect interactions that affect α_{ij}. The shape of this expression clarifies the form in which indirect effects — captured by the values of $\mathbf{\Lambda} = (\Lambda_1, \Lambda_2, \ldots, \Lambda_n)$ — feed into the optimality condition of the planner. The challenge consists in determining the exact shape of these indirect effects as a function of the network. It turns out we can obtain a recursive formulation for these constants in the spirit of other well known vertex similarity measures such as Katz-Bonacich or PageRank.[23] This is the content of Proposition 1.

Proposition 1. *The constrained-efficient risk sharing agreement for any network defined by G is characterized by a set of transfers given by,*

$$\alpha_{ij} = g_{ij}\left(M_j(\Psi,\mathbf{G}) - \frac{\Psi}{1+\Psi d_i}\sum_k g_{ik}M_k(\Psi,\mathbf{G})\right) \tag{9}$$

where $M_i(\Psi,\mathbf{G})$ corresponds to i's Weighted Even-Path Centrality (WEPC) defined recursively as,

$$M_i(\Psi,\mathbf{G}) = \frac{1}{d_i}\left(1 + \sum_{l,k} g_{ik}g_{kl}\frac{\Psi}{1+\Psi d_k}M_l(\Psi,\mathbf{G})\right) \tag{10}$$

Proof. See Appendix. □

Proposition 1 characterizes the full set of shares, $\mathbf{A}(\Psi,\mathbf{G})$, that defines the interior solution to the planner problem for any given network. As discussed above, the solution depends on the parameter Ψ and on the positions of each household in the network. The form in which the network defines the efficient sharing rule has to do with interactions among neighbors of neighbors (in other words, among

[21] Conversely, when controlling the variance becomes more important to the planner than covariance effects, the term Ψ decreases and the planner sets all of $j's$ shares much closer to each other, as demanded by the Λ_j term in (7)

[22] This is shown in the proof of Proposition 1.

[23] See, for instance, Leich, Holme and Newman (2006) for a theoretical account of vertex similarity in networks.

households located two links apart). To gain some intuition, recall the network interaction terms in equation (7): the shares going to household i are substitutes. This implies that households two links apart (with a common neighbor, say, i) interact directly as shown in equation (7). But indirect effects play a crucial role here as well. To see this, notice that these two households not only interact through their transfer to i, but also exchange resources with other partners, and these other relations affect what i receives from them, given their constraints that $\sum_i \alpha_{ij} = 1$. This is the main message behind equation (8). As a result, each household connected to i not only interacts directly with each other as in (7), but they also interact indirectly with others' sharing partners. The recursive definition of network centrality in Proposition 1 reflects these arguments: it says that $i's$ centrality depends on the centralities of $i's$ neighbors' neighbors (i.e. those households two links apart). Finally, Proposition 1 says that the sharing rule between any two households depends positively on the sender's measure, and negatively on the sum of measures of the receiver's neighborhood. It is sometimes helpful to think of this tradeoff as capturing the extent to which the sender's indirect interactions in the network cannot be accessed by any other of the receiver's partners.

The previous discussion argues that households at distance two interact directly, but also that households at distance four, six, eight etc.. interact indirectly. With this in mind, I show that the centrality measure captures all these direct and indirect effects across the network. The crucial element in this setting (following the previous arguments) consists of a network statistic that aggregates all *even-length paths* for every household, weighted in some particular way. In other words, Proposition 2 solves through the recursive definition of Proposition 1 in order to provide a reinterpretation of network centrality that clarifies the previous discussion.

Proposition 2. *Household $i's$ WEPC measure corresponds to a weighted sum of all* <u>*even-length paths*</u> *starting from i,*

$$M_i(\Psi, \mathbf{G}) = \frac{1}{d_i} + \sum_{q\in\mathbb{N}} \sum_{j\in N} \sum_{\pi_{ij}\in\Pi_{ij}^{2q}} W(\pi_{ij}) \tag{11}$$

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

$$W(\pi_{ij}) = \frac{1}{d_{i_0}} \frac{\Psi}{1+\Psi d_{i_1}} \frac{1}{d_{i_2}} \frac{\Psi}{1+\Psi d_{i_3}} \cdots \frac{1}{d_{i_q}} \tag{12}$$

Proof. See Appendix. □

Recursive measures like the one in equation (10) are common in the networks literature. These can be usually expressed similarly as the sum of all weighted paths starting from some household. I refer the reader to Section 8.3 for a more detailed and technical discussion of these graph measures and how the WEPC relates to them. In the context of the present discussion, however, it is interesting to note that contrary to other similar measures that weight all paths of a certain length equally, the current measure elicits path-specific weights, as described in equation (12). These weights reveal once more how the constraints (which weight all shares evenly as one over the degree, $\frac{1}{d_i}$) are used to connect long chains of interactions, in which two households interact via a common neighbor i through the term $\frac{\Psi}{1+\Psi d_i}$, as shown in equation (8).

In the next section I work through some properties of the sharing rule, and I argue the type of effects we expect in the estimation of standard tests for full insurance for some simple networks. In a way, the reader might want to think of this section as a more complete version of section 2, now that the sharing rule has been defined.

3.3 Comparative Statics and Implications for the Risk Sharing Test

Section 2 argues, like other recent papers, the importance of specifying the appropriate risk sharing group for each household when running the risk sharing tests. However, it also makes the case that coefficients to own income can generally be different from zero and that these coefficients vary across households, so that full risk sharing is rejected even at the appropriate sharing level; If networks are sufficiently symmetric, though, then households pass the Townsend (1994) test whenever local sharing groups are correctly controlled for. In this section I complement these arguments by working through two simple network structures using the sharing rule developed in the previous section.

First I show how, under symmetric structures, the sharing rule indeed boils down to a simple intuitive sharing behavior that predicts the standard efficiency results, when controlling for aggregate income of the relevant, household-specific sharing group. The most symmetric structure is the regular network — this is a network where every household is connected to k identical households, so all households are in identical positions. Following the arguments of section 2, we write household consumption in the form of risk sharing test

$$c_{it} = (\alpha_{ii} - \alpha_{ij})\, y_{it} + \alpha_{ij}\bar{y}_{it} + \epsilon_{it}$$

where $\bar{y}_{it}$ represents aggregate income of $i's$ local sharing group (i.e. $i's$ neighborhood). It is clear that the coefficient to $i's$ own income is zero (as in the Townsend tests) if $\alpha_{ii} = \alpha_{ij}$. This is true of regular networks.

Proposition 3. *The constrained-efficient sharing arrangement for any regular network with a common degree equal to k corresponds to the equal sharing rule defined as,*

$$\alpha^{\star}_{ij} = g_{ij}\frac{1}{k} \tag{13}$$

As a result, the first-best allocation is obtained for complete networks.

Proof. See Appendix. □

Of course the regular network is a very rare and extreme form of symmetry. In reality, social networks are far less structured and will therefore predict widely different transfers for different households. In these other cases, we expect instead that $\alpha_{ii} \neq \alpha_{ij}$ and therefore that the coefficients to own income will be positive and different across households. To take the most extreme example, consider the star network. In this network, one household is connected to all other households, that are otherwise not connected to anyone. In this case, I show the sharing rule simplifies to a simple expression relating to the size of the network and the level of connectivity of each household.

Proposition 4. *The constrained-efficient sharing arrangement for a star network of size n is given by,*

$$\alpha_{ij}^{\star} = g_{ij} \left(\frac{1}{1 + \left(\frac{n}{2} + 1\right) \Psi} \left(\frac{1}{d_j} + \Psi \right) \right)$$

for all $j \neq i$.

Proof. See Appendix. □

It is easy to see from this expression that flows towards the center consumes a smaller fraction of own income than other households. In fact, the star represents an extreme situation in which the center can very quickly be left to consume none of its own income. [24] Because the center mediates among a great number of other households, it benefits from diversification so long as incomes are sufficiently uncorrelated. Therefore, high centrality translates to lower consumption variance, and therefore the distribution of coefficients to own income across the population obtains a particular interpretation in terms of sharing opportunities — something that is not available in the standard risk sharing tests of Townsend (1994).

Finally, the value of Ψ can also provide drastically different predictions on the type of sharing behavior. Notice that if Ψ tends to zero — for instance for i.i.d. variables ($\rho = 0$) with small ratio $\frac{\mu^2}{\sigma^2}$ — then the optimality conditions imply that a household shares equally with all its neighbors (i.e. that $\alpha_{ij} = g_{ij}\frac{1}{d_j}$).[25] This is not surprising: extremely low values of Ψ represent situations where only minimizing aggregate volatility of uncorrelated earnings is important; this is accomplished by maintaining all shares as equal as possible.[26] This result poses a challenge in identifying the sharing behavior from data. Indeed if the true value of Ψ is low, it might be impossible to distinguish between sharing behavior described generally in Proposition 1 and a simpler, heuristic behavior such as equal sharing; both prescriptions should perform well as statistical models. In Section 3.4 I explore an alternative theoretical prediction of Proposition 1 in order to strengthen the belief that (9), and not (13), appropriately describes the sharing patterns of the Tsimane' communities.

On the other hand, as incomes correlate strongly across households the constrained-efficient sharing rule trades off diversification opportunities. In these situations, where Ψ is greater than zero (possibly

[24] More precisely, for $n > 4$ the center of the star consumes none of its own income if $\Psi \geq \frac{2}{n(n-4)}$ (for $n < 4$ interior solutions exist for all values of Ψ). This implies that as n increases, the space of parameters that guarantees an interior solution decreases.

[25] To see this notice that equation 7 in this case implies $\alpha_{ij} = \Lambda_j$ for all j, meaning that any two households connected to j receive the same fraction of $j's$ endowment. As a result, it must be that $\alpha_{ij} = \frac{1}{d_j}$. This represents equal-sharing.

[26] This is a classical iso-parametric problem of minimizing squares. Notice that for $\Psi \to 0$, the parameter σ^2 dominates over μ^2 and ρ so the planner problem is reduced to minimizing $\sum_{i,j} g_{ij}\alpha_{ij}^2$.

much greater), the planner's previous tendency to equate shares will lead to a large loss of surplus as strong correlation effects hike up consumption volatility. Now, the incentives move in the opposite direction and reigning in covariances is the primary concern; this is done by keeping all of $i's$ incoming shares as different as possible. In particular, I show below that at this other extreme— that is as Ψ tends to infinity— a household's net shares with each of its partners tend to zero.

Proposition 5. *As Ψ grows, the net exchange for any two households falls. In the limit, we have that*

$$\lim_{\Psi \to \infty} |\alpha_{ij} - \alpha_{ji}| = 0$$

for all $i \neq j$

Proof. See Appendix. □

Large values of Ψ correspond to situations of negligible net bilateral exchanges. This implies that household's in-shares correspond to its out-shares, that all households consume a convex combination of their partners' incomes.

4 Background and Data

In this paper I develop an alternative empirical specification to test risk-haring behavior by fitting (network) constrained-efficient exchanges across pairs of households, and recovering unexplained dependency between income and consumption. To do this I use panel data collected by a team of anthropologists from the years 2004 to 2009 in the small-scale, hunter-gatherer economies of the Tsimane' in the Bolivian Amazon. Before describing the data set in more detail I provide a quick description of the Tsimane' social structure, their economy, and general patterns of exchange (see Hooper (2011) for a much more thorough investigation into the economic life-cycle of the Tsimane').

4.1 The Tsimane' Indigenous Communities

The Tsimane' are an indigenous population of about 10 to 20, 000 individuals, residing in the Beni Department in lowland Bolivia. Tsimane' settlements are located primarily along the Maniqui and Quiquibey rivers, their tributaries and nearby forests. The Tsimane' organize primarily around a subsistence economy based on hunting, fishing, and slash-and-burn agricultural production of rice, sweet manioc (or yucca), plantain, and maize. According to Hooper (2011), most families maintain between 1 and 6 fields at one time (an average of 2.9 fields per family) that range in size from 0.1 to 2 hectares (an average of 0.56 hectares per field). While some of this production —- primarily, though not exclusively, rice — is sold to outside nearby markets in San Borja, still, around 95% of Tsimane' subsistence consumption rests on own production and exchanges across families.[27]

[27] See Martin, Melanie et al. (2012)

The Tsimane' social structure is primarily kin-oriented. Closely related nuclear families often reside together in small residential clusters, engaged in high levels of cooperative labor, common and shared meals, and other forms of resource pooling; bilateral exchanges of food across households account for the majority of this form of risk sharing.

The important gains that come from sharing uneven returns to productive effort are not foreign to the Tsimane'. The exchange of food across households forms a significant chunk of economic activity. In previous work on the Tsimane', Hooper (2011) shows clear evidence of reciprocity between families, and across types of goods, suggesting an interest in both attenuating risk and exploiting gains from specialization (see tables 5.1, 5.2 and 5.3 in Hooper (2011)). Around, 99% of the Tsimane' population engage in some form of food sharing at some point in the sample, and only 3.2% form separate trading groups of less than four households. From total production, an average of 5% is sold to outside and the rest is either consumed by the producing family, or exchanged with others. On average, 66% of a household's production is exchanged with other families, while 31% of a family's consumption consists of food received from other households. Genetic relatedness and the age of the household head interact as decisive attributes in determining the patterns of caloric exchanges. Hooper finds that while age alone does not seem to explain transfers, it nonetheless exhibits strong patterns of exchange between closely related families, not between unrelated ones. In terms of relatedness, it alone forms a very good predictor of food sharing. As an example, for two families with 40-year-old parents and zero net meat production, for example, the effect of a 0.1 increase in relatedness on the gross number of meat calories shared from one family to the other is 33.3 calories per day (Hooper, 2011).

4.2 The Data

The data comes from field work by a group of anthropologists at the Tsimane Health and Life History Project.[28] A series of field interviews were conducted from the years 2004 to 2009 on 250 families (1245 individuals) residing in 11 different Tsimane' villages. The villages are grouped into four separate regions: "downstream", "forest", "tributary", and "ton'tumsi". Figure 2.1 in Hooper provides a breakdown of the different sample periods and sizes. Each family was interviewed an average of 45.5 times (SD = 20.4), yielding a mean of 92.8 sample days per individual (SD= 40.0). Households were surveyed on average twice per week.

The surveys collected information on how many hours each family member spent laboring in subsistence activities during the preceding two days. These include hunting, fishing, and agricultural work. Quantities of edible products were recorded, and, for each product, interviewees were asked which members of the nuclear family, and which other community members had consumed portions of the product in prepared meals, or had received portions as raw gifts. Families were also asked whether they had received any gifts of food from other households.

For each product, the raw mass in kilograms was calculated from reported quantities based on mean

[28] Paul Hooper (Emory), Hillard Kaplan (UNM), and Michael Gurven (UCSB)

mass measurements derived from field guides and previous research with the Tsimane and other South American foragers. A product's total caloric value was based on estimates of mean dietary calories (assimilated by a human consumer after processing) per kilogram (Hooper, 2011).

In all sample communities, a detailed census was established that provided information on each individual's sex, birth year, and biological parents and grandparents.[29] Consanguineous and affinal relationships between individuals residing in the same community were calculated on the basis of shared genetic ancestry and marriage. Distance between households was constructed, when possible, from GPS data. (See Hooper, 2011) for a detailed account of all data collection procedures).

Together, this information constitutes an unbalanced panel of pairwise calorie exchanges at a frequency that is perhaps too high for this type of analysis. I therefore aggregate the data at a quarterly level. After discarding some pathological cases, I am left with 243 households in 8 different communities, sampled irregularly over 20 quarters, out of which an average household is sampled for 4.5 quarters (not always consecutive).[30]

4.3 Constructing Networks

As in most empirical studies of social networks, I confront the usual questions regarding how to define the appropriate underlying (and unobserved) social structures. As explained above, the unit of analysis is considered to be a nuclear family; I also refer to these as households. Although finer, within-household data on exchanges is available, all evidence suggests that these intra-family flows operate efficiently as completely connected (or in any case very dense) networks, so that for the purposes of this analysis they are best considered as distinct economic units.[31]

I test the model on three types of networks; each is accompanied with its own set of problems and advantages. The Trade Network establishes a link between two households if ever the two engage in any food sharing. This method of constructing links by "revealed preferences" of course fails to account for additional connections that could exist, but are otherwise not used. One could worry about endogeneity issues coming from this type of network. There are two things to say on this matter: First, the constrained-efficient exchanges I solve for are interior solutions to the planner problem, so the model speaks only to situations where all available links are utilized for some amount of food sharing, no matter how small. In other words, the model makes no predictions about which connections should be used, so taking observed trade as a link is not a huge problem. Secondly, It seems reasonable to assume that if two households share no calories throughout the entire sample then some social cost exists that impedes said relationship.

[29] Where incomplete, these census data were supplemented with data from demographic interviews described in Gurven et al. (2007). Adult parents and their co-resident dependents (i.e. offspring and adopted dependents) were classified together as nuclear families. Body mass in kilograms, assessed using an electronic standing scale, was available from yearly physical exams conducted by Bolivian physicians and research assistants for 1198 individuals in the sample (96%).

[30] I discard 5 households for which there is no reliable data. They appear to produce nothing and receive nothing throughout the entire sample. I drop another two households for which no reliable data on hours worked exists.

[31] See for instance, Hooper (2011).

It is customary in these communities for households to split upon marriage while remaining in the same village.[32] In order to account for this type of network modifications I also construct a dynamic version of the Trade Network. This network constructs links if, at every quarter, households are observed exchanging calories. For this particular network, then, the various centralities computed, and therefore the theoretical predictions on bilateral exchanges, are time dependent. While it might be unreasonable to assume in general that underlying social connections should change often, I show in the next section that in fact these networks show remarkable persistence over time, and in particular that the vast majority of links persist once they appear.

Finally, networks are also built using kinship data. To refrain from the endogeneity issues above I make no judgement on the "appropriate" level of kinship that determines the presence of a link. Instead, I construct links between households that share any level of genetic relatedness. This method poses its own set of concerns, not least of which has to do with missing genetic information for a number of households; this restricts the network artificially. Moreover, while kinship appears repeatedly in sociological work as a crucial determinant of social ties for a wide array of contexts, the Tsimane's population exhibit a disproportionately large share of endogamy[33]. This implies, as I describe in the next section, that while kinship networks might very well capture the appropriate dimension upon which social interaction develops, the level of connectivity might exceed any reasonable, underlying structure that determines insurance.

4.4 Descriptive and Network Statistics

Table 1 provides some descriptive statistics for various demographic and economic household attributes. Some of these represent cross sectional variability in different measures of household wealth, specified in Bolivianos (the national currency) and split into animal (livestock), traditional (non-mechanized productive tools) and modern (technological goods). These are used at times, together with demographic variables such as family size, age and marital status, to control for household-specific economic attributes.[34] More importantly, the average income and consumption variables represent the caloric production and flow data that are used extensively to test the model predictions on bilateral exchanges. These variables are longitudinal, specifying, for every household and every available date, the hours spent in each productive activity, the calories obtained as production, and the ensuing flows of those calories to nearby families. The data set also provides information on hours spent in three general productive activities: Agriculture, Fishing, and Hunting; Leisure is therefore defined as the number of hours in the past two days not spent in any of these activities.

Next I present network statistics for each of the kinship and trade networks that appear visually

[32]Hooper (2011) documents the creation and destruction of households in the Tsimane' context somewhere between 5 to 10% of households.

[33]See Hooper (2011)

[34]I mostly use household fixed effects to control for time-invariant household specific attributes such as these. However, when estimating the model's prediction on bilateral flows, edge-specific intercepts are unfeasible due to limited observations. In these cases I use a battery of controls such as those in Table 1, and others.

in figures 1 and 2. I show this information per village since I consider the villages as eight separate network structures. As described above, the kinship network is far denser, exhibiting larger cluster and closeness measures and much lower diameters across most villages. By way of comparing these two networks more rigorously, I also calculate the share of edges of each network that are observed in the other one. I find that about 71% of the connections observed in the data occur between households with some degree of genetic relatedness, while only 35% of households with genetic affiliation actually share food. [35] Finally, I present measures of persistence for the Updated Trade network to show that while it might be worthwhile to allow for certain network changes that come from the movement across villages or the creation of new families, networks are fairly stable over time.

In tables 5 to 8 I provide some preliminary evidence to support the assumption of no intermediation. Recall that this assumption implies that household consumption is a linear function of neighbors' incomes. As a result, two households that share a common neighbor will both consume a fraction of that income, and their consumption will be correlated. Households that are farther from each other, however, will not share any neighbors and their consumption will only co-move as per the underlying covariance across incomes. I therefore estimate consumption of each household against the aggregate consumption of those other households with which it shares a common neighbor (i.e. household in the set N_i^2), and against the aggregate consumption of households with which it does not share a common neighbor (i.e. households not in the set N_i^2). The results indicate that households farther away cannot explain consumption, once we control for aggregate income fluctuations.

5 Empirical Analysis

The theory above provides a number of predictions that can be tested directly against data on bilateral exchanges within networks. In this section I first run the standard test of full risk sharing and I find that full insurance is rejected in the Tsimane' data set. I then structurally estimate the sharing rule prescribed by the theory as given in equation (9); I find that the constrained-efficient prediction above appropriately describe the type of bilateral exchanges we observe across households. I also show that the model can retrieve the observed deviation from full insurance by estimating the risk sharing test on predicted consumption data. Finally, I test the model's implications on households' heterogeneous response to own income fluctuations. I find that the variation in household's coefficients to income follows the general pattern described by the model.

[35] I also perform a second measure of network comparison known as the Hamming distance, which measures the number of edges that need to be substituted to turn one network into another. We can see in Table 4that across most villages, we need to substitute about a third of all available dyads to move from the trade to the kinship network— in a couple of villages about half of that share is required.

5.1 Test of Full Risk Sharing

In order to assess how well the data conforms with the model, I explore a number of distinct predictions from the theory. Before I do this, however, I first perform the classical risk sharing test of Mace (1991), Cochrane (1991), and Townsend (1994), by running regressions of the form

$$\Delta log\,(c_{it}) = \beta_1 \Delta log\,(y_{it}) + \beta_2 \Delta log(X_{it}) + \tau_{vt} + \epsilon_{it} \tag{14}$$

where $\Delta log\,(c_{it})$ and $\Delta log\,(y_{it})$ stand for household consumption and income growth rates respectively, and τ_{vt} represents village-time fixed effects that capture uninsurable aggregate shocks that hit village v at time t. First differencing controls for any idiosyncratic time-invariant characteristic correlated with consumption; I also run some specifications in logs, in which case I add household-level fixed effects instead. Finally, X_{it} captures any other factors that could affect the optimal allocation of consumption and should be controlled for. In particular, for some specifications I control for household leisure over time, which affects consumption if preferences are non-separable and the planner cannot freely transfer leisure across households (Cochrane, 1991). In other specifications, I instead use leisure as an instrument to control for attenuation bias that might come from measurement error in the income variable. Leisure is a suitable instrument as it is undoubtedly correlated with income — households that spend more hours hunting, fishing, or harvesting will collect higher income, all else equal — but, because leisure is a separate survey item, measurement error in leisure arguably does not correlate with error in income (I follow Schulhofer-Wohl (2011) in this approach). All variables are expressed in adult-equivalent terms: I divide by a measure of a household's average adult caloric intake developed for the Tsimane' data set by Hooper (2011); it estimates caloric consumption across gender and age levels, and weights each household's demographic composition accordingly.[36] Standard errors are clustered at the household level.

Recall that we cannot reject the hypothesis of full risk sharing for values of $\beta_1 = 0$ and $\beta_2 = 1$. As shown in Table 9, I reject full insurance across all specifications. Coefficients on own income are about $\beta_1 = 0.35$ and statistically significant at the 1% level. Leisure is negative associated with consumption, as expected, but remains non-significant. Controlling for non-separabilities in income and consumption does not change the log estimates and lowers the growth rate estimates only by 0.005. The Instrumental Variables estimator controls for attenuation bias and therefore provides slightly higher estimates both for logs and growth rates; the difference, however, is quite small. All in all, I find a considerable correlation between consumption and own-income, consistent with previous studies in similar settings. Although the magnitude of this association varies across studies, a value of 0.35 falls well within the expected range. For instance, Munshi and Rosenzweig (2009) estimate values between 0.17 and 0.26 for Indian data, while Cochrane (1991) finds values between 0.1 and 0.2 in the PSID; Kinnan (2014) finds values ranging from 0.07 to 0.3 for Thai data, depending on the type of estimation.[37] Overall,

[36] See table 2.2 in Hooper (2011) for a more detailed explanation of this adult consumption measure

[37] Saidi (2015) finds that the magnitude of departure from efficiency is smaller than mine in the Tsimane' communities.

the results square fairly well with the literature and unequivocally reject full insurance. The theory provides new ways to think of partial insurance within a network context and help us understand the type of behavior that exists when we reject full insurance. To bring the main theoretical predictions to data, I first estimate the income process; I then fit the sharing rule to data.

5.2 Estimating the Income Process

Before taking my model to data, it is necessary to obtain an estimate for the income process. In this section I develop estimates of $\Psi = \frac{\mu^2+\rho}{\rho-\sigma^2}$ from panel data on the income processes of households. I carry out two different estimation procedures that deliver similar results. I first perform a very simple non parametric approach that uses basic moment estimators. I assume that income is described by two transitory shocks, an aggregate and an idiosyncratic one,

$$y_{it} = \kappa_t + \epsilon_{it}$$

with $\kappa_t \sim iid\left(\mu, \sigma_\kappa^2\right)$ and $\epsilon_{it} \sim iid\left(0, \sigma_\epsilon^2\right)$. I take the cross sectional average of income as an estimator of the aggregate shock, $\hat{\kappa_t} = \frac{1}{n}\sum_i y_{it}$, and I calculate the variability of this estimator to obtain an estimate of its variance $\hat{\sigma}_\kappa^2 = var(\hat{\kappa}_t)$. I obtain the variability in the residuals, $y_{it} - \hat{\kappa_t} = \hat{\epsilon}_{it}$, as an estimator of $\hat{\sigma}_\epsilon^2$. Finally, I compute the mean μ, the variance σ^2, and the common covariance term ρ, as follows: $\mu = \frac{1}{T}\sum_t \hat{k}_t$, $\sigma^2 = \hat{\sigma}_\kappa^2 + \hat{\sigma}_\epsilon^2$, and $\rho = \hat{\sigma}_\kappa^2$. These values deliver an estimate of Ψ equal to 0.97 ± 0.2.

I also perform a more sophisticated estimation procedure, availing myself of a vast and well established literature on estimating earnings processes from data.[38] The estimation assumes a state space model for the income process. Income is assumed to follow an aggregate shock, a temporary shock, and a persistent shock.

$$y_{it} = \kappa_t + \eta_{it} + \nu_{it}$$
$$\eta_{i,t} = c + \gamma\eta_{i,t-1} + \epsilon_{it}$$

This model is estimated by GMM. More information is given in Section 8.6, where I go over the estimation details and I show that I obtain values of Ψ close to those obtained with the more naive approach of the previous paragraph.

5.3 Structural Estimation of Network Flows

Having constructed networks and estimated the underlying income process, we are now ready to fit the model's sharing rules against the Tsimane's data set. To bring the model to data, recall the closed form expression for the constrained-efficient sharing rule as a linear function of global network measures

His data set comes from an entirely different survey corresponding to non-overlapping sets of Tsimane' villages. Moreover, he defines income from the sale of goods and labor as a separate survey item, whereas consumption here is income plus transfers, and therefore more tightly correlated.

[38]See, for instance, Lillard and Weiss (1979), MaCurdy (1982), Nakata & Tonetti (2015)

shown in equation (9),

$$\alpha_{ij}^{\star} = g_{ij}\left(M_j\left(\Psi, \mathbf{G}\right) - \frac{\Psi}{1+\Psi d_i}\sum_k g_{ik}M_k\left(\Psi, \mathbf{G}\right)\right) \tag{15}$$

and assume that shares are measured every period with additive error: observed bilateral shares are $\alpha_{ijt}^{obs} = \alpha_{ijt} + \epsilon_{ijt}$. Then the constrained-efficient sharing rule proposed in this paper implies the following relationship,

$$\alpha_{ijt}^{obs} = \beta_1 M_j + \beta_2 M_{N_i} + \epsilon_{ijt} \tag{16}$$

where $M_{N_i} = \frac{\Psi}{1-\Psi d_i}\sum_k g_{ik}M_k\left(\Psi, \mathbf{G}\right)$ aggregates the WEPC centrality measure across of all of $i's$ neighbors. Under such a specification the theory requires that $\beta_1 = 1$, and $\beta_2 = -1$. I control for village-level shocks by allowing for village-time specific intercepts and control for a number of household-specific attributes, such as household size, total wealth, marital status, and average age of the household heads.[39] I show in Table 10 that we obtain estimates statistically indistinguishable at the 5% level from $\beta_1 = 1$ and $\beta_2 = -1$ for the Kinship, Trade, and Updated Trade networks, using a value of $\Psi = 0.9$ as estimated in the previous section. Values for the Trade and Update Trade networks are statistically indistinguishable at the 1% level. Moreover, in Figures 6 to 8 I plot the two coefficients for a wide range of Ψ values to clarify that these results are fairly robust to possible estimation errors. Both coefficients are statistically different from zeros. We see that β_2 clearly takes on values close to -1 for values of Ψ away from zero and that for most values it is statistically indistinguishable from -1 at the 95% confidence level. Already in this first direct approach the model performs remarkably well at describing the observed relationship between network structure and household exchanges.

I also perform a more demanding test of the model. Ideally, I would like to estimate the sharing rule separately for each pair of households over time, and obtain distinct coefficients to each of the centrality measures in equation (15). The lack of sufficient longitudinal data, however, precludes this type of analysis. As an alternative, notice that the second term in equation (15) varies according to the degree of the receiving household. I exploit this variability by splitting the population by their degree, d_i, and redefining M_{N_i} in (16) as, $M_{N_i} = \sum_k g_{ik}M_k\left(\Psi, \mathbf{G}\right)$. In this case, a successful model would obtain a negative coefficient of $\beta_2\left(d_i\right) = -\frac{\Psi}{1+\Psi d_i}$ that increases with the degree. Figures 10 and 11 plot $\beta_2\left(d_i\right)$ against the degree of each group. I also plot the curve representing the theoretical prediction of $\frac{\Psi}{1+\Psi d_i}$. The positive relationship between this coefficient and the degree of the receiving household is clear. Moreover, the model and theoretical predictions move in a similar fashion and reasonably close to each other.

Although the model seems to fit the sharing data quite well, we might still worry that other network centrality measures could also predict similar results, undermining the model's predictive

[39] Because I assume networks to be undirected, any edge that only sustains unilateral exchanges is complemented by adding a flow equal to zero in the opposite direction for any period in which the household in question obtains positive production.

power. After all, it is well known that network centrality measures often correlate strongly. In 8.5 I discuss the statistical relationship between the centrality measure proposed here and several other familiar candidates in the network literature. I show that these other measures are not strongly correlated with the WEPC centrality that I propose, and more importantly that they fail to explain the patterns of exchange observed in the Tsimane' data.

5.4 Revisiting the Risk-Sharing Test

In this section I show that the positive coefficient on own income obtained in the test of full insurance of Section 5.1 can be interpreted as capturing the bilateral sharing arrangement proposed in this paper. To do this I run the income data through my model to obtain household consumption under the constrained-efficient arrangement that I propose. I then estimate the risk sharing test of equation (14) with predicted, rather than observed, consumption data. A successful theory of partial insurance would retrieve the same coefficients to own income as those observed in the data in section 5.1.

More concretely, I use the sharing rule of Proposition 1 to calculate the consumption level of each household in every period, given income, as,

$$\hat{c}_{it} = \sum_j g_{ij} \left(M_j - \frac{\Psi}{1 + \Psi d_i} \sum_k g_{ik} M_k \right) y_{jt}$$

where I have dropped the explicit dependency on $\mathbf{G}$ and Ψ for convenience. This equation defines household consumption using the proposed sharing rule of the model. Expected consumption data is used, in lieu of the actual consumption, to test whether the type of variability in household consumption behavior proposed by this model can replicate the departure from full efficiency observed in the data.

The results are presented in Tables 12 to 14 for the different networks being analyzed, and for the value of Ψ estimated in section 4.3. Under my model's predictions, and for the output data available for the Tsimane', the coefficient on own income corresponding to the Trade Network shown in Table 13 oscillates between 0.15 ± 0.023 for OLS to about 0.25 ± 0.06 for IV estimates. Compared to the value of about 0.35 ± 0.07 that we obtain in Table 9, the model seems to slightly underestimate the empirical loading on own income for this network structure. However, it is worth noting that these differences are not large, and that for the IV estimates the difference is statistically not significant. The results for the kinship network presented in Table 12 show estimates far too low to resemble the magnitude of departure from efficiency in Table 9; the specification in growth rates provides OLS estimates that are non significant, suggesting full insurance under network constraints. This is not particularly surprising given that, as mentioned above, kinship networks are excessively dense, so that the planner problem is far less constrained than in the other network structures.[40] On the other hand, the results for the Updated Trade network in Table 14 are statistically indistinguishable from the estimates in Table

[40]For this same reason this network performs worst in the structural estimation of bilateral exchanges of section 4.2 for the value of Ψ estimated from data.

9. Although these coefficients, again, lie below the values provided by data, the differences now are negligible— sometimes as small as 0.01 — and therefore are all statistically insignificant.

5.5 Underlying Heterogeneity in Consumption

The environment I describe not only accounts for the type of coefficients we obtain when we reject full insurance, but, more importantly, defines a complete distribution of these coefficients based on network measures. Moreover, as argued in Section 2, the size of these coefficients can provide information about the relative sharing opportunities of each household in certain environments. Indeed, households with lower coefficients to their own income process are far more central than others, and as a result obtain in general smoother consumption paths. In other words, while the previous section showed that the model can generate a common coefficient that reflects the observed departure from efficiency, the following estimation procedure implies that the theory also provides insight into the type of asymmetric insurance possibilities affecting households as a result of their social situation.

Recall that a household's share of its own income left for consumption can be described, for each i, as,

$$\alpha_{ii} = M_i\left(\Psi, \mathbf{G}\right) - \frac{\Psi}{1+\Psi d_i}\sum_k g_{ik} M_k\left(\Psi, \mathbf{G}\right) \tag{17}$$

Retrieving precisely these values from data would require estimating each household's theoretical consumption, as described in (17), independently. The short time dimension of the panel unfortunately prohibits this type of analysis. Moreover, the variation across these values is often small for "similar" nodes and would be difficult to extract from the inherent noise in data. Instead, I decide to rank the population according to (17) and then split the population into equally sized groups. I then estimate equation (14) separately for each group; This gives me enough variability both within and across groups to effectively measure the expected positive difference across successive groups. [41]

The results are shown in Figures 12 and 13. The positive trend across groups is evident for all networks, and is especially pronounced for the trade and updated trade networks. In all these cases, while any two consecutive groups might show little variation, the overall increase from the first to the last group is generally about 0.5, and in all but the kinship network the difference is statistically significant. In other words, the positive association between income and consumption found for the Tsimane' data set can be further decomposed into those households that, by the overall social arrangement, consume more or less of their own income ex-post.

[41] In order to allow for as much intergroup variability as possible, group size was kept as small as possible, while retaining enough observations to provide efficient standard errors. Average group size was 35 households per group, leading to a total of 7 groups.

6 Conclusion

Time and again, evidence collected from risk-sharing communities in the developing world has concluded that households in these type of arrangements are only partially insured against random fluctuations in income. In this paper I argue that these insurance mechanisms overwhelmingly perform below full efficiency precisely because networks of interactions are not completely connected. I show that if the underlying social structures are accounted for when deriving constrained-efficient exchanges, then observed trades across pairs of households is well described by the theory, and the distance from the Pareto frontier can be obtained.

I propose a constrained-efficient framework that relaxes a crucial assumption in the classical risk-sharing literature, which allows all households to trade with each other. Instead, I restrict the movement of goods along a given set of social relations and I derive a full analytical description of the exchanges between any two households as a function of their network position. I show that exchanges are determined by a global network measure that accumulates all direct and indirect interactions along the entire network. In other words, this theory endogenizes pairwise sharing behavior along any given network. The theory is useful in providing a rich description of the type of partial insurance we might expect if we believe network constraints are a relevant friction keeping communities below full efficiency. More importantly, it can be easily tested in a number of different settings, as long as income and network data is available. In this sense, it is capable of providing testable predictions at the pairwise level, generating much more detailed variation on the exchanges generating consumption streams.

I test the theory with data from Tsimane' indigenous communities in the Bolivian Amazon. I structurally estimate the constrained-efficient sharing rule against bilateral exchanges observed across Tsimane' households and find that the theory does a good job of fitting empirical sharing behavior. Moreover, predicted consumption profiles generate the type of inefficiencies observed in these communities, and other important implications on the distribution of insurance levels across different households are also observed in data. Overall, evidence from Tsimane' communities suggest that accounting for incomplete social structures goes a long way to explain the type of partial insurance mechanisms operating more broadly in village economies.

Of course a number of other elements have been proposed that surely form part of a full description of these complicated social arrangements. For instance, Ligon et al. (2000) have studied the presence of limited commitment in these informal exchanges and argue that incentive constraints under dynamic contracts indeed lead to partial insurance similar to that first observed by Townsend (1994) for Indian villages. More recently, Schulhofer-Wohl (2015) and Mazzocco and Saini (2012) have stressed that heterogeneous preferences might lead one to overestimate the failure of full insurance. I believe these views and the one I propose here are complementary and together build a richer story of informal insurance. Indeed my model refrains from considering these and many other interesting dimensions, and I try and stay as close as possible to the classical setup of risk-sharing proposed by Mace (1991), Cochrane (1991), and Townsend (1994), while at the same time allowing me to engage directly with a general network structure.

A network description of exchanges like the one proposed here holds great promise for identifying vulnerable households, or determining superior social arrangements. After all, one of the advantages of modeling social interactions explicitly in this context is that it provides a great deal of heterogeneity on consumption volatility and inequality both across households and networks. It would be interesting to know, for instance, which arrangements perform better than others, and whether we can generally identify households that, if removed, would most affect the sharing opportunities of the entire community. Indeed network models like this one have already answered these type of questions in other settings, such as Ballester et al. (2007) which do a similar exercise for criminal networks. Although I provide some tentative results on the ranking of households by consumption volatility in section 8.4, the ordering is only partial and I am currently working on new results. The challenge here, with respect to Ballester et al. (2007), has to do with the complicated weighting scheme for paths that emerges in this setting, and which is absent in the Bonacich centrality or other similar recursive network measures. Indeed a lot of the existing tools to make progress on this front utilize the convenient geometric weighting of Bonacich, but I am working on a recursive formulation of WEPC that would allow me to make progress nonetheless.

Another ambitious proposal that emerges from this analysis seeks to structurally estimate the underlying social structure. In other words, if we agree the model performs well in this context, and might therefore be a good proxy for the type of sharing behavior of the Tsimane', then an exciting step forward would utilize the theory's predictions in order to structurally back out an estimate of the true underlying structure. This approach is not without its own set of challenges, not least of which is that the model requires inputing an entire network described by an n^2 dimensional object. Extracting this from data is not simple. However, there have been some recent developments by Manresa (2015) on estimating the structure of interactions from panel data using a pooled lasso estimator that might be very useful. If we can frame the spillover effects in an amenable way, it might be possible to identify the most likely structure from within the class of sparse networks.

Perhaps the most promising step forward involves more general results on the welfare implications from my theory. Indeed, a theoretical result that relates network-based heterogeneity in consumption behavior to more general welfare implications would provide a clear economic interpretation to the coefficients of empirical risk sharing tests. In other words, beyond rejecting or not full insurance, the theory could allow data to speak more clearly on the distribution of welfare across the population when full insurance is rejected. The distribution of households' response to income shocks that this paper predicts could then be mapped directly to a normative implication on welfare. It would form an important contribution, and would come full circle towards a new interpretation of empirical risk sharing test.

7 Tables and Figures

This section presents all tables and figures in the order in which they appear in the text. Some additional tables and figures can be found in the Online Appendix.

	n	mean	sd	median	min	max
HH Size	245	5.22	3.04	5.00	1.00	16.00
Mean Age Head of HH	245	36.96	16.33	34.00	14.50	86.00
No. of Dependents	245	3.12	2.74	2.00	0.00	13.00
% Rice Sold on Market	184	32.03	25.54	29.23	0.00	95.00
Animal Wealth (Bolivianos)	199	1478.28	3500.81	455.00	0.00	28250.00
Traditional Wealth (Bolivianos)	199	1179.94	1078.11	743.00	0.00	5917.50
Modern Wealth (Bolivianos)	199	3582.21	2356.61	3348.64	184.68	10726.22
Total Wealth (Bolivianos)	199	6259.17	5089.62	5403.40	363.20	35582.94
Avg. Income (Calories)	244	925.63	694.70	750.30	0.00	3896.01
Avg. Consumption (Calories)	246	888.02	694.77	707.66	46.67	4527.67
Avg. Out Flow (Calories)	242	353.53	536.24	174.18	0.00	4365.31
Avg. In Flow (Calories)	246	286.84	415.54	167.88	0.00	3441.35
Avg. Leisure (Hours)	244	46.22	1.48	46.61	35.97	48.00

Table 1: Household Summary Statistics: *Variables expressed in adult-equivalent terms. Averages taken over periods where data is available*

	n	Edges	Avg.Degree	Diameter	Density	Cluster	Avg.Between	Avg.Closeness
1	27	73	5.407	6	0.193	0.596	0.026	0.066
2	38	121	6.368	5	0.163	0.441	0.055	0.344
3	11	45	8.182	2	0.682	0.676	0.042	0.753
4	20	46	4.600	6	0.219	0.484	0.058	0.115
5	13	51	7.846	3	0.560	0.624	0.054	0.642
6	27	189	14.000	3	0.500	0.627	0.024	0.645
7	46	122	5.304	10	0.113	0.357	0.064	0.205
8	65	320	9.846	5	0.149	0.315	0.022	0.422

Table 2: Network Statistics Per Village: Trade Network

	n	Edges	Avg.Degree	Diameter	Density	Cluster	Avg.Between	Avg.Closeness
1	26	194	14.923	5	0.287	0.687	0.043	0.229
2	38	292	15.368	4	0.202	0.806	0.017	0.077
3	11	111	20.182	2	0.917	0.940	0.010	0.928
4	20	110	11.000	5	0.275	0.810	0.029	0.092
5	13	121	18.615	3	0.716	0.871	0.030	0.779
6	26	210	16.154	4	0.311	0.781	0.025	0.155
7	44	768	34.909	4	0.397	0.691	0.016	0.609
8	64	1594	49.812	5	0.389	0.810	0.013	0.581

Table 3: Network Statistics Per Village: Kinship Network

	Hamming Distance	Normalized Hamming Distance
1	100	0.285
2	116	0.165
3	18	0.327
4	29	0.153
5	30	0.385
6	128	0.365
7	320	0.309
8	804	0.387

Table 4: Hamming Distance per Village between Trade and Kinship Networks

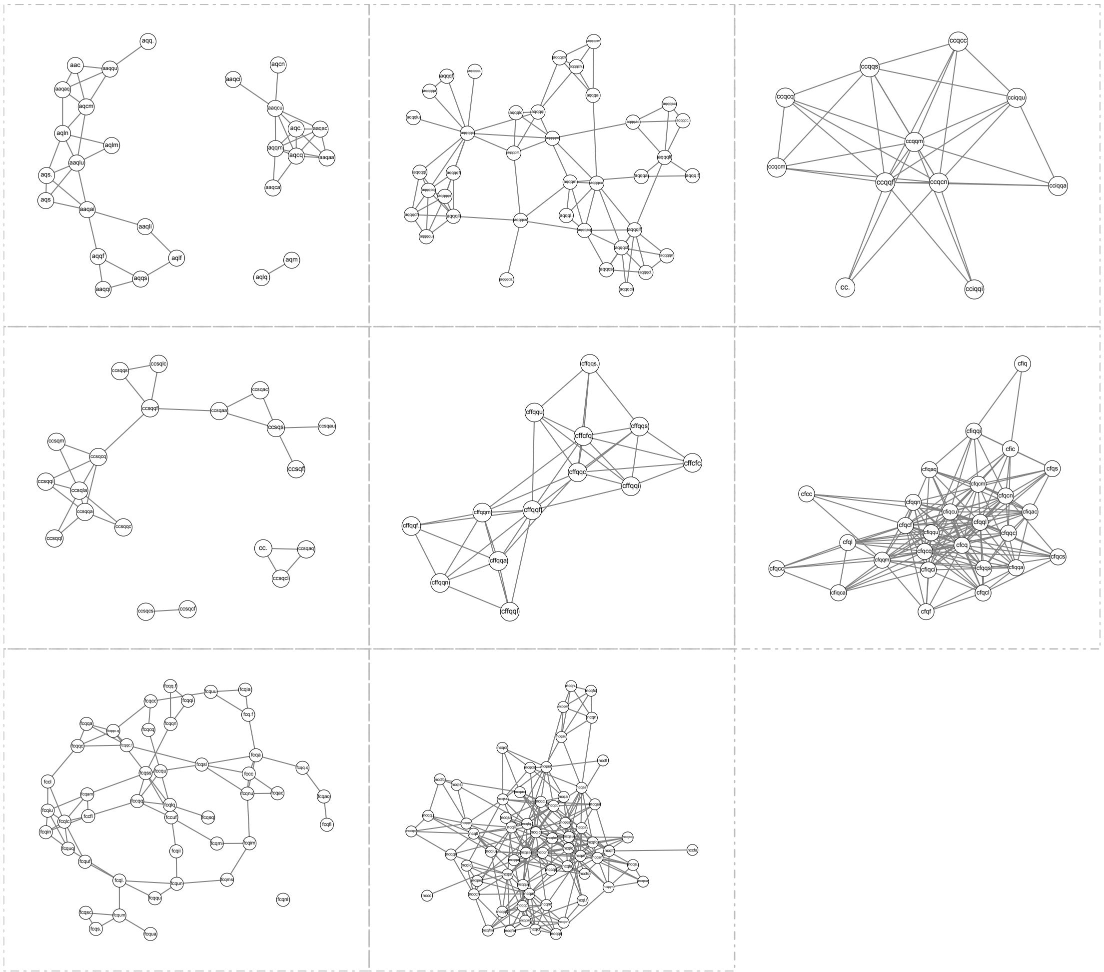

Figure 4: Trade Network: Link exists if households exchange food at any point in the sample.

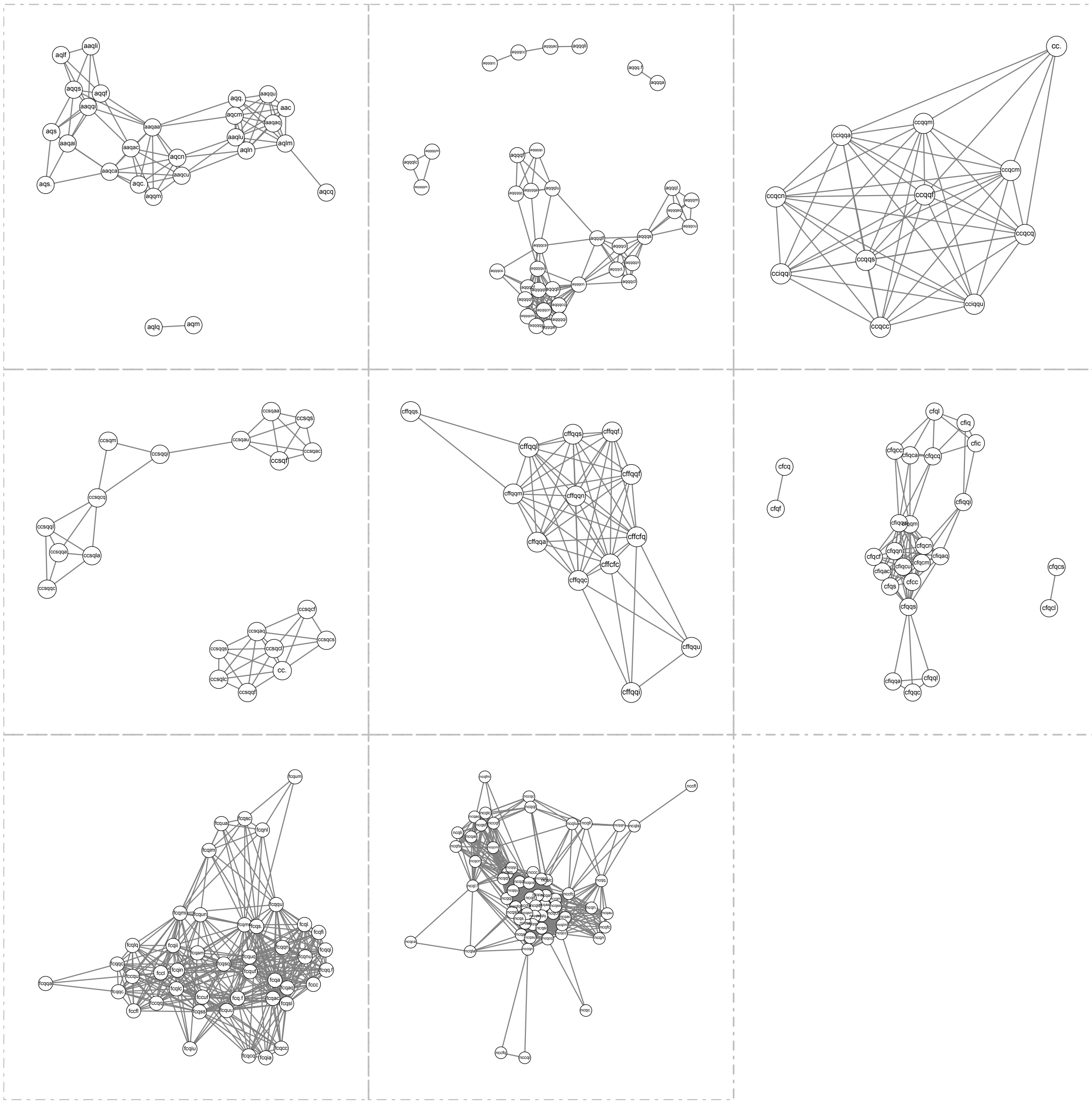

Figure 5: Kinship Network: Link exists if Mean Genetic Relation is above 0

Table 5: Local Correlations: Trade Network

	Dependent variable:					
	log(consumption)			Δ log(consumption)		
	No IV		Instrument: Leisure	No IV		Instrument: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(income)	0.362***	0.360***	0.375***			
	(0.039)	(0.040)	(0.039)			
log(leisure)		−0.344				
		(2.452)				
log(consumption $\in N_i^2$)	0.298***	0.299***	0.295***			
	(0.107)	(0.108)	(0.106)			
log(consumption $\notin N_i^2$)	0.214	0.215	0.215			
	(0.145)	(0.145)	(0.143)			
Δ log(income)				0.378***	0.376***	0.396***
				(0.036)	(0.037)	(0.037)
Δ log(leisure)					−0.463	
					(1.933)	
Δ log(consumption $\in N_i^2$)				0.117	0.117	0.117
				(0.103)	(0.102)	(0.102)
Δ log(consumption $\notin N_i^2$)				0.096	0.095	0.099
				(0.114)	(0.115)	(0.111)
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	922	922	922	692	692	692
R^2	0.470	0.470	0.470	0.485	0.485	0.485
Adjusted R^2	0.350	0.349	0.349	0.463	0.462	0.463

Note: *p<0.1; **p<0.05; ***p<0.01

Variables constructed as log(1+x) to admit zeros in cons. and inc. data

Values in parentheses are standard errors clustered at the household level.

Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

Table 6: Local Correlations: Kinship Network

	Dependent variable:					
	log(consumption)			Δ log(consumption)		
	No IV		Instrument: Leisure	No IV		Instrument: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + inc)	0.361***	0.365***	0.326***			
	(0.040)	(0.042)	(0.040)			
log(leisure)		0.873				
		(2.041)				
log(consumption $\in N_i^2$)	−0.439***	−0.447***	−0.444***			
	(0.165)	(0.168)	(0.167)			
log(consumption $\notin N_i^2$)	−0.114	−0.112	−0.120			
	(0.108)	(0.106)	(0.108)			
Δ log(income)				0.363***	0.363***	0.367***
				(0.038)	(0.039)	(0.038)
Δ log(leisure)					−0.077	
					(1.632)	
Δ log(consumption $\in N_i^2$)				−0.437***	−0.437***	−0.436***
				(0.155)	(0.156)	(0.155)
Δ log(consumption $\notin N_i^2$)				−0.160	−0.160	−0.159
				(0.108)	(0.107)	(0.108)
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	878	878	878	669	669	669
R^2	0.477	0.478	0.476	0.506	0.506	0.506
Adjusted R^2	0.358	0.357	0.357	0.481	0.480	0.481

Note: *p<0.1; **p<0.05; ***p<0.01
Variables constructed as log(1+x) to admit zeros in cons. and inc. data
Values in parentheses are standard errors clustered at the household level.
Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

Table 7: Local Correlations: Trade Network

	Dependent variable:					
	log(consumption)			Δ log(comsumption)		
	No IV		Instrument: Leisure	No IV		Instrument: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(income)	0.353***	0.350***	0.378***			
	(0.039)	(0.040)	(0.039)			
log(leisure)		−0.659				
		(2.450)				
log(consumption $\in N_i$)	0.235***	0.236***	0.232***			
	(0.077)	(0.077)	(0.076)			
log(consumption $\in N_i^2 \cap \notin N_i$)	−0.077	−0.077	−0.081			
	(0.064)	(0.065)	(0.064)			
log(consumption $\notin N_i^2$)	0.155	0.155	0.157			
	(0.144)	(0.144)	(0.139)			
Δ log(income)				0.369***	0.369***	0.368***
				(0.036)	(0.038)	(0.036)
Δ log(leisure)					0.013	
					(1.972)	
Δ log(consumption $\in N_i$)				0.190**	0.190**	0.190**
				(0.080)	(0.081)	(0.080)
Δ log(consumption $\in N_i^2 \cap \notin N_i$)				−0.128*	−0.128*	−0.128*
				(0.069)	(0.069)	(0.069)
Δ log(consumption $\notin N_i^2$)				0.071	0.071	0.071
				(0.108)	(0.108)	(0.108)
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	914	914	914	682	682	682
R^2	0.471	0.471	0.470	0.501	0.501	0.501
Adjusted R^2	0.349	0.349	0.349	0.476	0.476	0.476

Note: *p<0.1; **p<0.05; ***p<0.01
Variables constructed as log(1+x) to admit zeros in cons. and inc. data
Values in parentheses are standard errors clustered at the household level.
Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

Table 8: Local Correlations: Kinship Network (Kinship)

	Dependent variable:					
	log(consumption)			Δ log(consumption)		
	No IV		Instrument: Leisure	No IV		Instrument: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(income)	0.357***	0.360***	0.328***			
	(0.040)	(0.042)	(0.040)			
log(leisure)		0.733				
		(1.938)				
log(consumption $\in N_i$)	−0.047	−0.053	−0.049			
	(0.107)	(0.102)	(0.106)			
log(consumption $\in N_i^2 \cap \notin N_i$)	−0.375***	−0.378***	−0.388***			
	(0.120)	(0.120)	(0.118)			
log(consumption $\notin N_i^2$)	−0.098	−0.096	−0.103			
	(0.114)	(0.112)	(0.113)			
Δ log(income)				0.363***	0.363***	0.362***
				(0.038)	(0.039)	(0.038)
Δ log(leisure)					0.012	
					(1.688)	
Δ log(consumption $\in N_i$)				−0.099	−0.100	−0.100
				(0.105)	(0.104)	(0.105)
Δ log(consumption $\in N_i^2 \cap \notin N_i$)				−0.268**	−0.268**	−0.268**
				(0.105)	(0.105)	(0.105)
Δ log(consumption $\notin N_i^2$)				−0.151	−0.151	−0.151
				(0.106)	(0.105)	(0.106)
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	874	874	874	665	665	665
R^2	0.476	0.477	0.476	0.503	0.503	0.503
Adjusted R^2	0.356	0.356	0.355	0.477	0.476	0.477

Note: *p<0.1; **p<0.05; ***p<0.01
Variables constructed as log(1+x) to admit zeros in cons. and inc. data
Values in parentheses are standard errors clustered at the household level.
Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

Table 9: Full Risk Sharing Test

	Dependent variable:					
	log(consumption)			Δ log(consumption)		
	OLS		IV: Leisure	OLS		IV: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + inc)	0.355***	0.355***	0.357***			
	(0.036)	(0.020)	(0.036)			
log(leisure)		−0.043				
		(1.260)				
Δ log(income)				0.360***	0.355***	0.402***
				(0.034)	(0.035)	(0.036)
Δ log(leisure)					−1.070	
					(1.635)	
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	1,129	1,129	1,129	854	854	854
R^2	0.452	0.452	0.452	0.469	0.469	0.467
Adjusted R^2	0.339	0.339	0.339	0.449	0.449	0.448

Note: *p<0.1; **p<0.05; ***p<0.01

Variables constructed as log(1+x) to admit zeros in cons. and inc. data

Values in parentheses are standard errors clustered at the household level.

Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

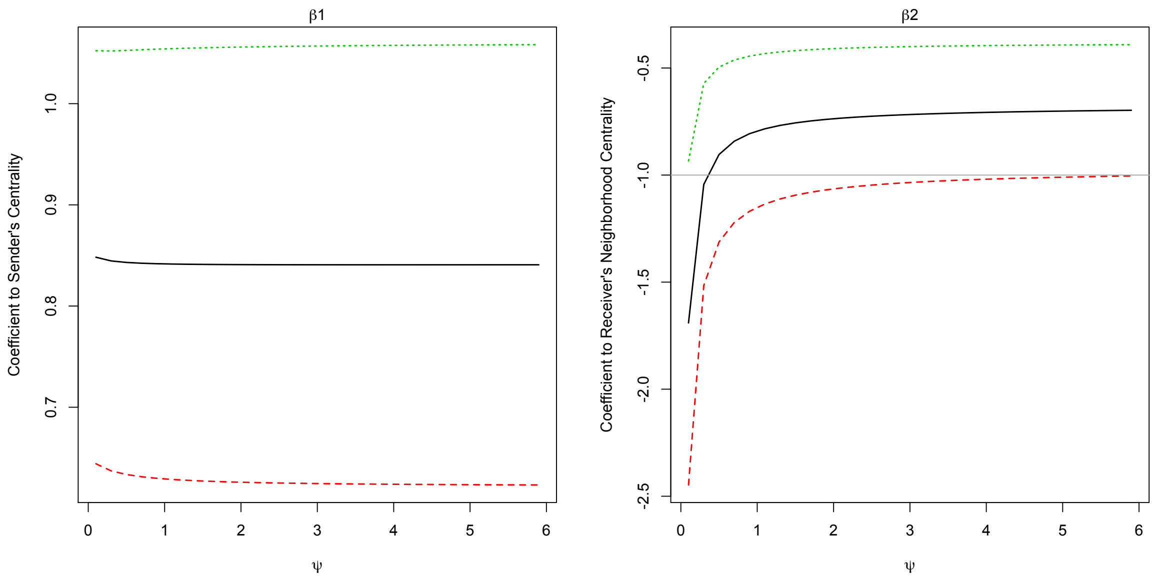

Figure 6: Coefficients to Network Centralities in Regression of Edge-Level Exchanges: Trade Network (Households younger than 4 0)

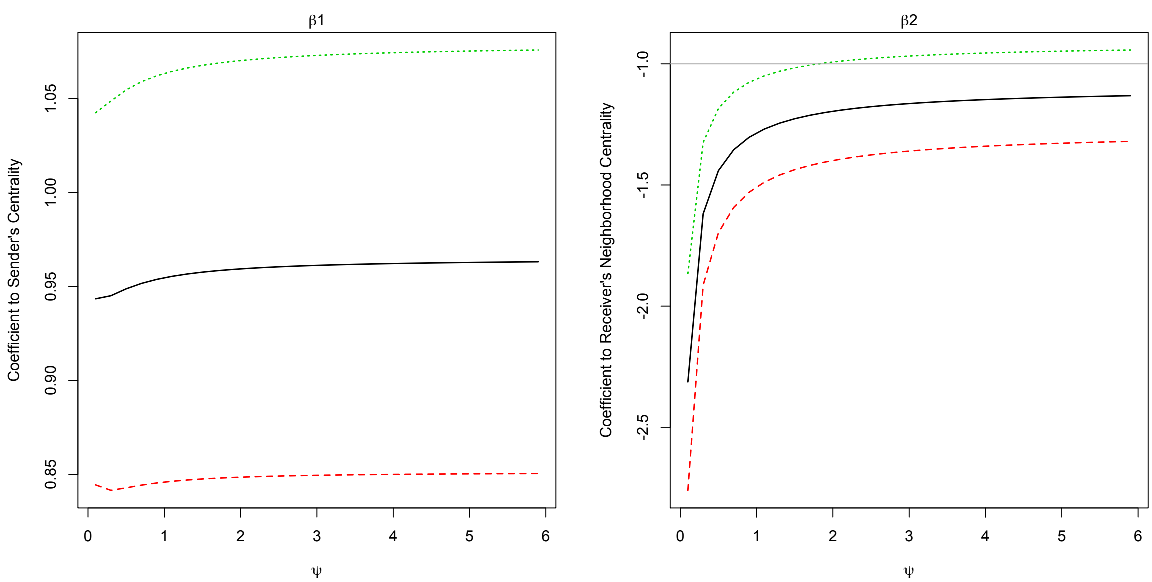

Figure 7: Coefficients to Network Centralities in Regression of Edge-Level Exchanges: Kinship Network

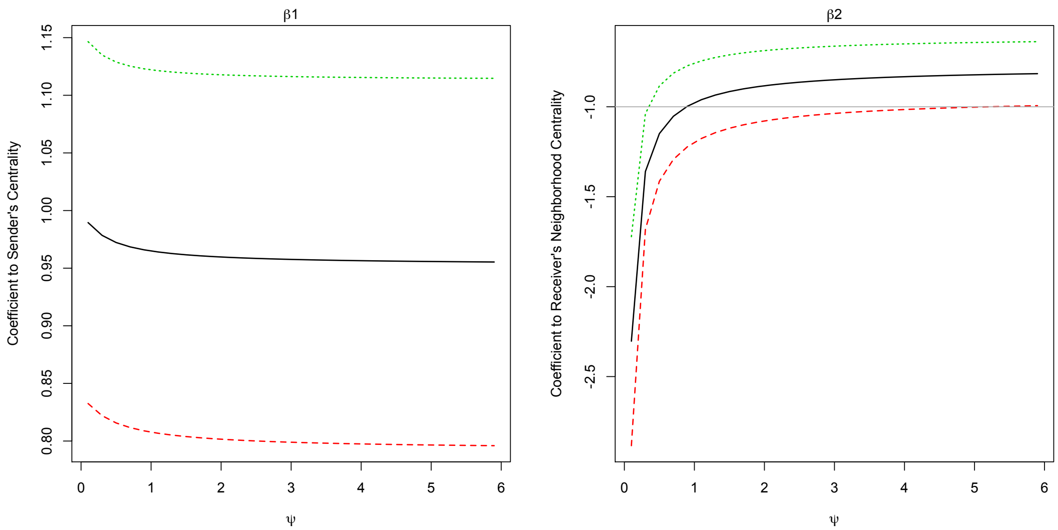

Figure 8: Coefficients to Network Centralities in Regression of Edge-Level Exchanges: Updated Trade Network

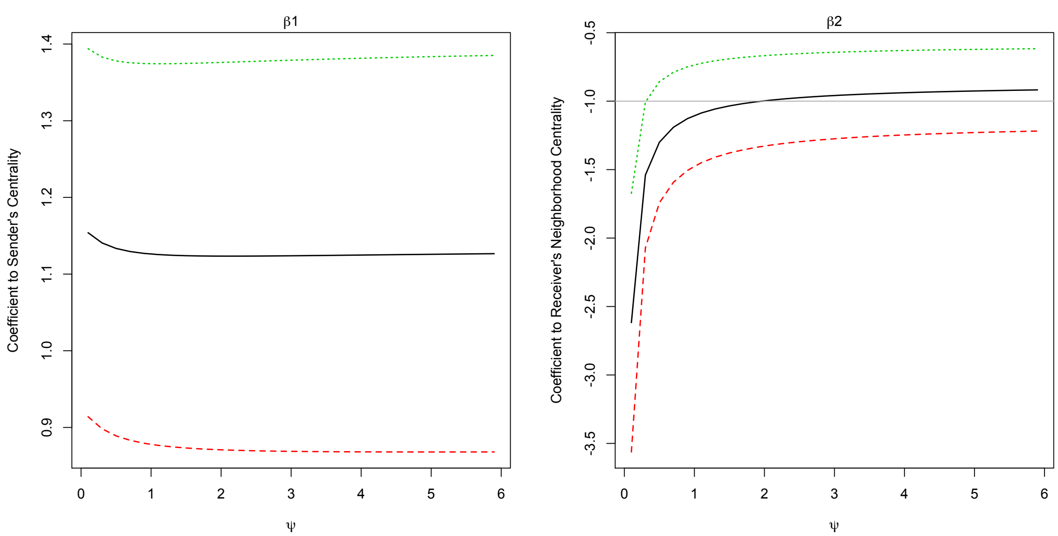

Figure 9: Coefficients to Network Centralities in Regression of Edge-Level Exchanges: Updated Trade Network (Households younger than 4 0)

Table 10: Regression of Edge-Level Exchanges on Predicted Sharing Rule ($\Psi = 0.9$)

	Dependent variable:		
	α_{ij} (i.e. share from i to j)		
	Kinship Network	Trade Network	Updated Trade Network
	(1)	(2)	(3)
$M_j(\mathbf{G}, \mathbf{\Psi})$	0.878***	0.841***	1.109***
	(0.044)	(0.111)	(0.093)
$M_{N_i}(\mathbf{G}, \mathbf{\Psi})$	−1.324***	−0.697***	−0.946***
	(0.113)	(0.156)	(0.133)
Village-Time Fixed Effects	Y	Y	Y
Observations	11,943	2,059	1,730
Adjusted R^2	0.092	0.180	0.202

Note: *p<0.1; **p<0.05; ***p<0.01
Values in parentheses are standard errors clustered at the household level.

Table 11: Regression of Edge-Level Exchanges on Alternative Local Measure

	Dependent variable:		
	α_{ij} (i.e. share from i to j)		
	Trade Network	Kinship Network	Updated Trade Network
	(1)	(2)	(3)
$\frac{\lvert N_i^2(\mathbf{G})\rvert}{\sum_i g_{ij}\lvert N_i^2(\mathbf{G})\rvert}$	0.096**	0.167***	−0.007
	(0.049)	(0.037)	(0.054)
Village-Time Fixed Effects	Y	Y	Y
Observations	5,586	11,931	2,742
Adjusted R^2	0.057	0.043	0.080

Note: *p<0.1; **p<0.05; ***p<0.01
Values in parentheses are standard errors clustered at the household level.

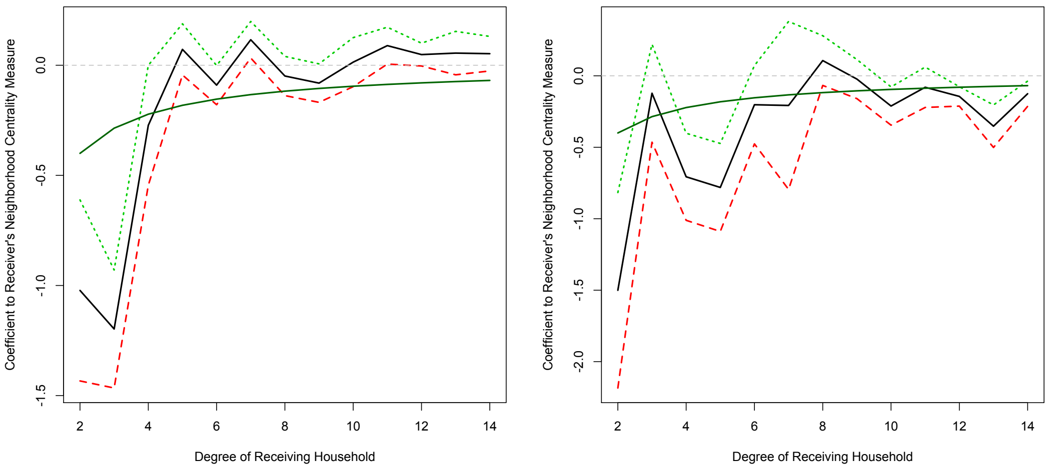

Figure 10: Coefficient β_2 as a function of Receiver's degree. Panel A: Trade Network. Panel B: Kinship Network

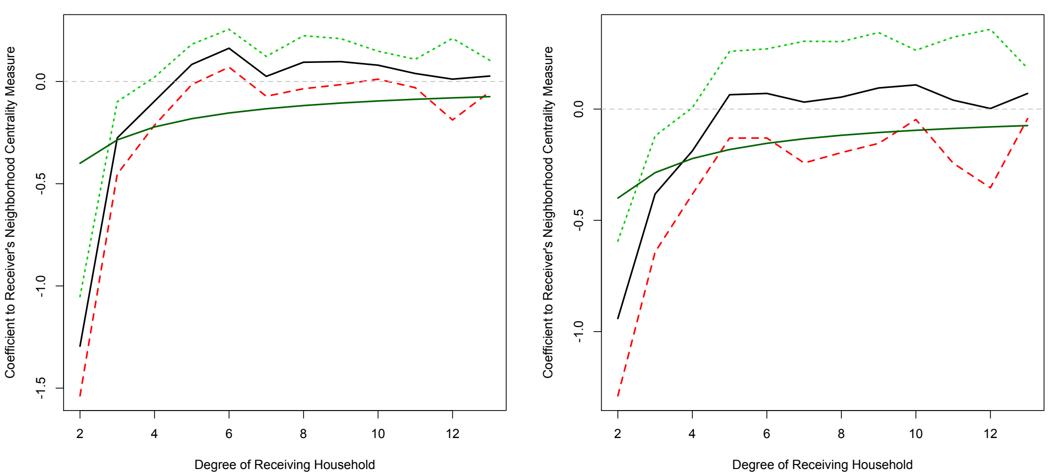

Figure 11: Coefficient β_2 as a function of Receiver's degree. Panel A: Updated Trade Network. Panel B: Updated Trade Network (Households younger than 40)

Table 12: Full Risk Sharing Test for Model Consumption Data: Kinship Network

	Dependent variable:					
	log(Predicted Consumption)			Δ log(Predicted Consumption)		
	OLS		IV: Leisure	OLS		IV: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(income)	0.062***	0.056***	0.158***			
	(0.016)	(0.016)	(0.021)			
log(leisure)		−1.282*				
		(0.711)				
Δ log(income)				0.026	0.017	0.141***
				(0.018)	(0.018)	(0.026)
Δ log(leisure)					−1.928**	
					(0.773)	
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	1,100	1,100	1,100	817	817	817
R^2	0.406	0.408	0.372	0.334	0.340	0.279
Adjusted R^2	0.302	0.303	0.277	0.319	0.324	0.266

Note: *p<0.1; **p<0.05; ***p<0.01
Variables constructed as log(1+x) to admit zeros in cons. and inc. data
Values in parentheses are standard errors clustered at the household level.
Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

Table 13: Full Risk Sharing Test for Model Consumption Data: Trade Network

	Dependent variable:					
	log(Predicted Consumption)			Δ log(Predicted Consumption)		
	OLS		IV: Leisure	OLS		IV: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + inc)	0.154***	0.148***	0.248***			
	(0.022)	(0.023)	(0.026)			
log(leisure)		−1.260*				
		(0.718)				
Δ log(income)				0.133***	0.124***	0.243***
				(0.025)	(0.025)	(0.031)
Δ log(leisure)					−1.851**	
					(0.883)	
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	1,096	1,096	1,096	812	812	812
R^2	0.398	0.399	0.380	0.343	0.347	0.317
Adjusted R^2	0.295	0.296	0.282	0.327	0.330	0.303

Note: *p<0.1; **p<0.05; ***p<0.01
Variables constructed as log(1+x) to admit zeros in cons. and inc. data
Values in parentheses are standard errors clustered at the household level.
Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

Table 14: Full Risk Sharing Test for Model Consumption Data: Updated Trade Network

	Dependent variable:					
	log(Predicted Consumption)			Δ log(Predicted Consumption)		
	OLS		IV: Leisure	OLS		IV: Leisure
	(1)	(2)	(3)	(4)	(5)	(6)
log(1 + inc)	0.316***	0.312***	0.397***			
	(0.042)	(0.043)	(0.046)			
log(leisure)		−0.948				
		(1.152)				
Δ log(income)				0.286***	0.278***	0.392***
				(0.043)	(0.044)	(0.049)
Δ log(leisure)					−1.607	
					(1.485)	
Household Fixed Effects	Y	Y	Y	N	N	N
Village-Time Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	1,086	1,086	1,086	800	800	800
R^2	0.401	0.402	0.396	0.360	0.362	0.351
Adjusted R^2	0.297	0.298	0.294	0.344	0.345	0.336

Note: *p<0.1; **p<0.05; ***p<0.01
Variables constructed as log(1+x) to admit zeros in cons. and inc. data
Values in parentheses are standard errors clustered at the household level.
Instrumental Variables estimation uses Balestra-Varadharajan-Krishnakumar's transformation.

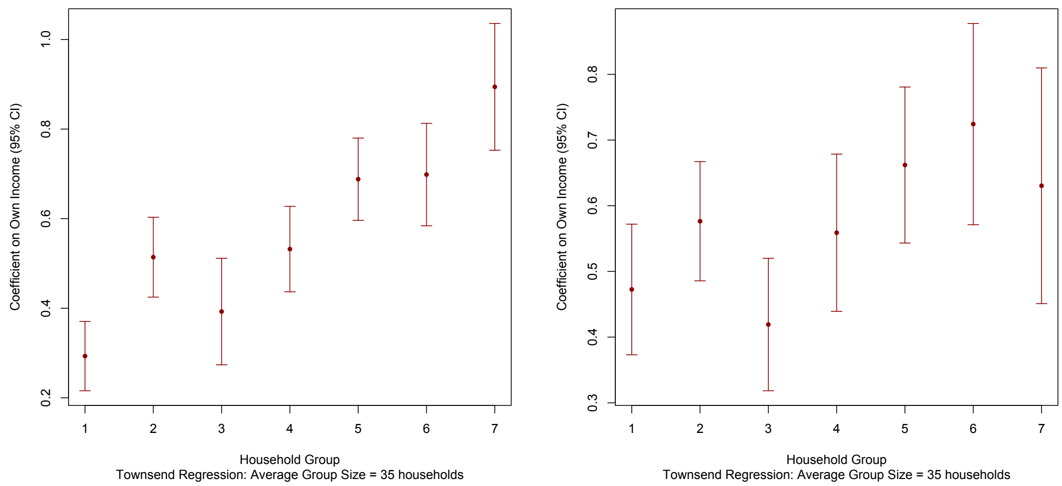

Figure 12: Coefficients and Confidence Intervals for Equation 14 Partitioning Population according to Centrality Measure. Panel A: Trade Network. Panel B: Kinship Network

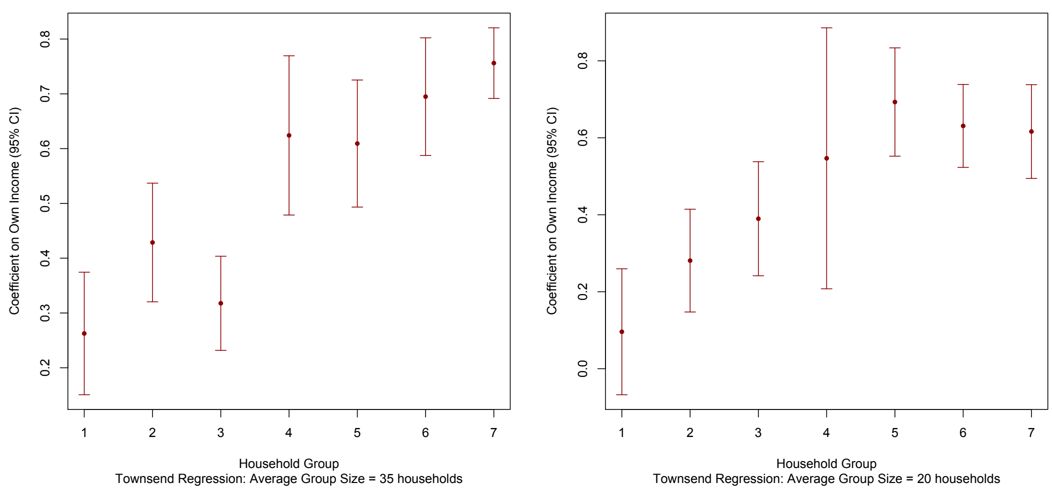

Figure 13: Coefficients and Confidence Intervals for Equation 14 Partitioning Population according to Centrality Measure. Panel A: Updated Trade Network. Panel B: Updated Trade Network (Age < 40)

8 Additional Results

8.1 Contingent Sharing Rules

Consider the set of contingent sharing rules that maximize welfare for the simple economy in Figure 1. Shares from i to j now depend on the state of the world ω; intuitively, the distribution of income in each state will determine the sharing opportunities. For instance, if household 1 obtains an income $y_1(\omega)$ larger than $y_2(\omega)$ and $y_3(\omega)$ then funds can readily be redistributed so that the efficient condition of equation (1) holds for that particular state ω. However, it is easy to see that this will not be possible for all states: for instance if $y_1(\omega) < \bar{y}(\omega) - y_2(\omega)$ then the income of household 1 is not large enough to transfer the required resources to household 2.

Intuitively, because sharing groups are local, household consumption is bounded above by equation (3). A good way to think of this setting more generally is by imagining that the lack of intermediation essentially sets capacity constraints on what each node can transfer to its neighbors. This type of environment is explored in Ambrus et al. (2014) in the context of credit constraints on a network with exogenous link values. Here, rather than limiting the transfer across nodes by

As a result, only partial insurance is possible and the ratio of marginal utilities is not constant across all states. As we have just seen, the ratio is constant only for a subset of states where the income of the intermediating household is sufficiently large. More generally we can define a set $\bar{\Omega}(\mathbf{G}) \subseteq \Omega$, such that for a given network $\mathbf{G}$ it provides the subset of Ω such that full insurance is possible. The previous discussion signals the inherent difficulty in isolating general network effects from particular income realizations for these type of contingent sharing rules. Not surprisingly, a large part of the literature on risk sharing networks have dealt with fixed (or non-contingent) sharing rules.

8.2 A Model with Network Intermediation

In this section I show how to extend the current setup to a allow for network intermediation. In particular, I relax the assumption of no-intermediation to a general case where households can access income from households at some distance k (the setup analyzed in the main text corresponds to the situation where $k = 1$). To simplify the arguments, let $k = 2$ in what follows. However, all the arguments below apply for all values of k . In this scenario, consumption by household i is a linear function, not only of incomes of neighbors (as before), but also of the income of neighbors' neighbors (i.e. those households two links away from i) as follows,

$$c_i(\omega) = \sum_{jk} g_{ik} g_{kj} \alpha_{ik} \alpha_{kj} y_j(\omega)$$

It is easy to see that the relationship between consumption and income is still defined by my model's predictions as given in Proposition 1, but where the primitive of the model now is not the original network $\mathbf{G}$, but rather a new network $\tilde{\mathbf{G}}$ that is built from $\mathbf{G}$ by connecting all households that are two links apart. Indeed, while the exact trade is hard to describe analytically as a function of the true network $\mathbf{G}$, it is nonetheless very easy to describe as a function of the network $\tilde{\mathbf{G}}$: it depends on the sharing rule of Proposition1. In other words, the theoretical predictions of my model allow for general descriptions of intermediation, and the general arguments on the implications for consumption behavior follow through. The only caveat is that it is now difficult to characterize the type of exchanges (in this case exchanges of exchanges) that will lead to an efficient solution, but the efficient solution, in terms of consumption responses to income shocks, can be described precisely by my model. Finally, notice that there is an upper limit on the amount of intermediation that generates a situation of partial insurance. Indeed, if k is larger than the minimum distance separating any two households, then intermediation is sufficiently large that full insurance is retrieved.

8.3 Discussion of Weighted Even Path Centrality

Proposition 1 is powerful because it provides a full description of efficient sharing-behavior under restricted bilateral exchanges for all possible social networks and distributional parameters. As a concrete prescription of network flows to be tested against data, it suffices, and, as we will see in section 3, performs reasonably well. In this section I describe the expression of the WEPC in equation (10) in more detail and discuss what it can tell us about the optimal network shares.

It turns out that recursive expressions like the one in (10) are found often in network analysis. These measures attempt to quantify associations between vertices based solely on the structure of connections. For instance, in their well-known work on strategic complementarities in networks, Ballester et al. (2007) show that equilibrium actions depend on a similar recursive measure known as Bonacich Centrality. More recently, Banerjee et al. (2012) have sought to identify individuals in the network that are best placed to diffuse information on microcredit opportunities in India. They find that participation is higher if those first informed have higher eigenvector centrality. [42] It is a matter of fact that global network measures such as these always appear in situations with entangled interactions along a set of connections. All of these measures can be expressed generically as

$$B_i = c + \gamma \sum_k g_{ik} B_k \tag{18}$$

for some constant c and with $|\gamma| < 1$. This expression essentially says that $i's$ measure depends linearly on the sum of measures that are connected to i. Let us distinguish two crucial differences with respect to the WEPC measure defined in (10). First of all, notice that equation (10) does not sum over all measures that i is connected to, but instead sums over all measures that $i's$ partners are connected to. In other words, the WEPC is defined recursively at distance two, not one. This is not entirely rare in network analysis and in fact appears in some work on vertex similarity by Jeh and Widom (2002). It

[42] 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.

has also appeared in newer page-ranking algorithms, such as the HITS algorithm.[43] Secondly, notice that, in contrast to equation (18), the WEPC does not weight all neighboring measures with a common parameter γ. Instead, the measures at distance two are 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. Moreover, notice that the particular weight has the familiar form of equation (8) that captures the full extent of $l's$ local interactions with its partners. Remember that $M_i(\mathbf{G}, \Psi)$ captures all of $i's$ indirect interactions along the network. Following earlier discussions, these indirect interactions affect i only in so far as they alter the sharing behavior of those at distance two from i (with whom i actually interacts). It only seems natural, then, that $i's$ measure, $M_i(\mathbf{G}, \Psi)$, is defined recursively from the measures, $M_l(\mathbf{G}, \Psi)$, for all households l that that i directly interacts with (i.e. those that are at distance two from i). In other words, the exact shape in the recursive definition of $M_i(\mathbf{G}, \Psi)$ spells out quite clearly the preceding discussion on how indirect interactions appear in the local tradeoffs of each household. I show next that the recursiveness in (10) can be undone into an expression that holds a lot more meaning in terms of a household's network position.

An important property of expressions like (18) is that, for appropriate values of ϕ — particularly for $\gamma < \frac{1}{\nu_1}$ for ν_1 the largest eigenvalue of $\mathbf{G}$ — we can write them as,

$$\mathbf{S} = \mathbf{I} + \gamma\mathbf{G} + \gamma^2\mathbf{G}^{\mathbf{2}} + \ldots$$

In other words, all these recursive measures are often expressed as the sum of all paths starting from i, which can be written as $B_i = \sum_{j\in N}\sum_{q=1}^{\infty}\gamma^q g_{ij}^q$. This framework provides a much more natural way to think of the network statistic containing information on the importance of each household in the general social structure. After all, it measures a household's accessibility by aggregating all locations that can be reached by it at any length. Since we have just seen that the WEPC is a particular version of these general recursive measures, it should not be surprising that it too can be written as the accumulation of paths. Proposition 2 indeed shows that we can similarly think of the WEPC measure accumulating such paths, subject to the caveats discussed above: that only even paths are accumulated, and that the weighting scheme for each path is particular to that path. Technically, the current setting asks us to solve a modified version of these fixed points on a graph that looks like, $(\mathbf{D} - \mathbf{G}\Psi\mathbf{G})^{-1}$. I show in Proposition 2 that we can write this as,

$$\mathbf{D}^{-1}\left(\mathbf{I} + \left(\mathbf{A} + \mathbf{A}\mathbf{D}^{-1}\mathbf{A} + \mathbf{A}\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-\mathbf{1}}\mathbf{A} + \mathbf{A}\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-\mathbf{1}}\mathbf{A}\mathbf{D}^{-1}\mathbf{A} + \ldots\right)\mathbf{D}^{-1}\right)$$

Unlike Bonacich and other types of centralities that weight all paths of a certain length equally, in

[43]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.

this scenario each path elicits a specific set of weights, determined by the connectivity of each individual involved in that particular path. The type of weighting scheme in equation 2 can be thought of in terms of the accumulated local interactions mentioned above. Recall that indirect interactions only represent the concatenation of various direct interactions linked together by the network constraints. This can be gleaned from equation 2 where the weights $\frac{\Psi}{1+\Psi d_k}$ capture all the households in a given path engaged in direct interactions and the weighs $\frac{1}{d_i}$ capture the connecting household's constraint. This weighting scheme marks a crucial distinction vis-à-vis other measures, in that additional paths does not guarantee an increase in households' network measure.

8.4 Individual and Aggregate Volatility

In the context of bilateral risk-sharing in networks a natural concern seeks to distinguish amongst those households that, from their structural position in a broader social arrangement, obtain smoother consumption streams than others. Said differently, we can ask how a household's consumption variability relates to its location in the network. Proposition 1 defines transfers and therefore establishes, for every household, a particular linear combination of neighboring incomes that enter its consumption. The particularly complicated form of these transfers makes it difficult to obtain an intuitive translation from network position to consumption variance. In the following result I provide a first, partial attempt at ranking household's consumption variance from network characteristics; I am currently working on expanding this into a complete, intuitive ordering of variances on networks. Notice we can write the variance of consumption of household i as ,

$$var\left(c_i\right) = \left(\sigma^2 - \rho\right)\alpha_i'\alpha_i + \rho\mathbf{1}'\alpha_i\alpha_i'\mathbf{1}$$

where $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \ldots \alpha_{in})'$ is defined as in (9). We want to find the household $i^\star$ such that $var\left(c_{i^\star}\right) \geq var\left(c_i\right)$ for all $i \neq i^\star$. Using the expression for exchanges in (9) This next result allows us to rank variances when endowments are independent across households.

Proposition 6. *Let* $\mathbf{H}_i = \left(diag\left(\mathbf{G}_i\right) - \frac{\Psi}{1+\Psi d_i}\mathbf{G}_i'\mathbf{G}_i\right)$. *If* $\rho = 0$ *and* $\mathbf{H}_i^2 - \mathbf{H}_j^2$ *is positive semi-definite, then* $var\left(c_i\right) > var\left(c_j\right)$

Proof. See Appendix □

The rather technical form of this result precludes a straightforward interpretation on the distribution of consumption volatility. In any case, it provides a testable prediction on individual consumption volatility that is fairly quickly checked in data. I am in the process of extending this result and testing it on the Tsimane' data set.

Aggregate volatility of an entire village is perhaps even more important than distinguishing amongst individual variances. After all, policy considerations can emerge from a deeper understanding of what social arrangements are more conducive to better insurance opportunities. In this respect, we might want to know 1) what type of network is persistently less volatile than another, or 2) what individual, when removed, reduces volatility the most. After some manipulations, I obtain a "useful" form that should allow me to conclude something about which networks are prone to higher aggregate volatility. Specifically, I have that,

$$\sum_i var\left(c_i\right) \propto \mathbf{1}'\left(\mathbf{D}-\mathbf{G}\boldsymbol{\Psi}\mathbf{G}\right)^{-1}\left(\mathbf{D}-\mathbf{G}\boldsymbol{\Psi}\boldsymbol{\Gamma}\mathbf{G}\right)\left(\mathbf{D}-\mathbf{G}\boldsymbol{\Psi}\mathbf{G}\right)^{-1}\mathbf{1}$$

where $\boldsymbol{\Gamma}$ is a diagonal matrix similar to $\boldsymbol{\Psi}$, such that $\boldsymbol{\Gamma}_{ii} = \frac{2+\Psi d_i}{1+\Psi d_i}$. The familiar quadratic form, although a complicated function of the adjacency matrix, might conceal some useful properties that might allow me to answer these two questions.

8.5 Alternative Centrality Measures

As in most network papers that prescribe a centrality-based prediction on behavior, a natural concern is that in fact other similar measures might be as successful in explaining data, so that the predictions of the model are rendered mute. Most times this is dealt with by running a horse race against other measures and showing that the theory's predictions indeed outperform other measures. To do this, I first show that the WEPC centrality presented in this paper is only weakly correlated with other well known global network measures, such as Bonacich centrality or eigenvalue centrality. Correlation with Bonacich is about 0.32 for the Kinship Network and about 0.28 for the Trade Network. Correlation with eigenvalue centrality is about 0.46 for the Kinship Network and about 0.38 for the Trade Network. These values are not large, moreover if we substitute these measures for the WEPC in the expression for the sharing rule in equation (9), we obtain non-significant, and even negative, result. Of course, these values are not normalized, so they predict shares that fail to satisfy the constraint (i.e. outside the interval $[0,1]$ and/or don't sum to one).

A possible objection, therefore, may be that any other linear function of arbitrary network measures that both defines values in $[0,1]$ and satisfies the feasibility constraint, $\sum_j \alpha_{ij} = 1$, would deliver similar results; said differently, one could ask if (16) estimates nothing but a simple accounting identity of cross-claims on a network. Indeed, while there exist many such matrices $\mathbf{A}\left(\mathbf{G}\right)$ that satisfy the feasibility constraints, the estimation procedure could fail to distinguish amongst them, delivering "appropriate" fits to vastly different predictions. To test this I estimate a simple, intuitive alternative to equation (9) that only captures local node characteristics. Specifically I consider the possibility that the share of $j's$ endowment consumed by i is determined entirely by the size of $i's$ $k-neighborhood$ relative the total of all $k-neighborhoods$ of all of $j's$ neighbors. This measure captures $i's$ relative importance within $j's$ sphere of influence similar to (9), but under a reduced, local notion of importance. We can express this sharing behavior as,

$$\alpha_{ij} = \frac{\left|N_i^k\left(\mathbf{G}\right)\right|}{\sum_i g_{ij}\left|N_i^k\left(\mathbf{G}\right)\right|}$$

where $|A|$ denotes the cardinality of set A. Indeed it is not difficult to see that this first-order measure provides predictions within the unit interval and satisfies the budget constraint. If the structural estimation of the sharing rule only captures an accounting identity, the estimation of

$$\alpha_{ijt}^{obs} = \beta_1 L_{ij} + \epsilon_{ijt} \tag{19}$$

where $L_{ij} = \frac{|N_i(\mathbf{G})|}{\sum_i g_{ij}|N_i(\mathbf{G})|}$, should undoubtedly produce $\beta_1 = 1$. I show results for the $2-neighborhood$ in table 11, although similar results hold for all k values tested (all below 10). The results show that indeed β_1 is either not significant, or far from one. Instead, relative neighborhood size fails to correlate with the sharing behavior of the Tsimane' in any reasonable manner that would indicate that other local measures can do as good a job at describing pairwise exchanges.

8.6 Estimating the Income Process

Before estimating the income process I control for predictable components. Although the data set contains a number of time-invariant demographic statistics for each household, the only time-varying, household-specific attribute that predicts the level of income is hours worked. Therefore, I run the following first-stage regression of log household income on hours invested in productive activities, together with household and village-time fixed effects,

$$log\left(y_{i,t}\right) = h_{i,t} + \tau_{vt} + \delta_i + \epsilon_{it} \tag{20}$$

I choose to allow for household-specific intercepts rather than introducing a long, but still incomplete, list of household demographic traits. I obtain a residual income process for household i from (20) that I use as my unpredictable component of income in order to estimate the parameter Ψ.

The next step requires that we define a process for residual income,

$$\tilde{y}_{it} = \eta_{it} + \nu_{it}$$
$$\eta_{i,t} = c + \gamma\eta_{i,t-1} + \epsilon_{it}$$

where $\tilde{y}_{i,t}$ is the residual from a log income regression for an individual i at time t, η_{it} is the persistent component of income and is assumed to follow an AR(1) process, ν_{it} is the transitory component of income, ϵ_{it} is the shock to the persistent component of household income. Finally, $\upsilon_{it} \sim \left(0, \sigma_\upsilon^2\right)$, $\epsilon_{it} \sim \left(0, \sigma_\epsilon^2\right)$, $\eta_{i,0} \sim \left(0, \sigma_0^2\right)$ and are independent of each other for all i and t. The parameter vector to estimate is $\theta = \left(\gamma, c, \sigma_\epsilon^2, \sigma_\upsilon^2\right)$. Notice that we don't make any distributional assumptions on the error terms besides defining first and second moments.

Before estimating the vector θ I relate its elements to the parameter of interest, Ψ. Recall that $\Psi = \frac{\mu^2+\rho}{\sigma^2-\rho}$ where μ, σ^2, and ρ represented the mean, variance and common covariance term of the joint distribution of income across households. Given the description on residual income above we can conclude the following relationship between the parameters of θ and the parameters that form the value of Ψ:

$$\mu = \frac{c}{1-\gamma}$$
$$\sigma^2 = \sigma_\epsilon^2 \frac{1}{1-\gamma^2} + \sigma_\nu^2$$
$$\rho = \gamma \frac{c}{1-\gamma}$$

Computing the cross-sectional covariances between period t and period $t+k$ (for all t and k) produces a total of $\frac{T(T+1)}{2}$ distinct moment conditions that relate residual income and distributional parameters.[44] In particular, if we write down the moment $m_{tk}(\theta)$ between agents at time t and $t+k$ we have,

$$\begin{aligned} m_{tk}(\theta) &= \mathbb{E}\left[y_{i,t} \cdot y_{i,t+k}\right] \\ &= \mathbb{E}\left[\left(\eta_{it} + \nu_{it}\right)\left(\eta_{i,t+k} + \nu_{i.t+k}\right)\right] \\ &= \begin{cases} \mathbb{E}\left[\eta_{it}^2\right] + \sigma_\nu^2 & if\ k = 0 \\ \gamma^k \mathbb{E}\left[\eta_{it}^2\right] + \left(1-\gamma^k\right)\mathbb{E}\left[\eta_{it}\right]^2 & if\ k > 0 \end{cases} \end{aligned}$$

where

$$\mathbb{E}\left[\eta_{it}^2\right] = \sigma_\epsilon^2 \frac{1}{1-\gamma^2} + \mathbb{E}\left[\eta_{it}\right]^2$$

and $\mathbb{E}\left[\eta_{it}\right] = \frac{c}{1-\gamma}$ is the typical expression for the mean of an AR(1) process. The above expressions represent an over-identified system for θ, so moment conditions cannot be solved explicitly.[45] As usual in these cases, we look for the vector θ that minimizes the distance between theoretical moments and their empirical counterparts,

$$\hat{\theta} = \min_\theta \left(M(\theta) - \hat{M}\right)' \mathbf{W} \left(M(\theta) - \hat{M}\right)$$

where $M(\theta)$ and $\hat{M}$ stack the moment conditions and the sample analogs respectively, and where $\mathbf{W}$ is a weighting matrix. Following the general trend in the literature I take $\mathbf{W}$ as the identity matrix.[46] The non-linear GMM estimation delivers estimates of $\hat{\theta} = \left(\hat{\gamma}, \hat{c}, \hat{\sigma_\epsilon^2}, \hat{\sigma_v^2}\right) = \left(0.981, 0.00932, 0.521 * 10^{-5}, 2.018\right)$. This represents a negligible shock to the persistent component of income, and a strong persistence parameter. the variance to transitory shocks is about 2. This implies that household income can best be thought mostly of transitory shock with a small common intercept. Using the expressions above we find that $\mu = 0.93$, $\sigma^2 = 2.019$, and $\rho = 0.86$. Together this implies an estimated value of $\Psi = 1.474$. This value is close to other values found using more rudimentary estimates of simpler models in the

[44] Notice the matrix of moment conditions is symmetric (i.e. $m_{tk}(\theta) = m_{kt}(\theta)$) so we only calculate the lower triangular part of the matrix, consistent of $\frac{T(T+1)}{2}$ distinct terms.

[45] In unbalanced panels like this one, moreover, we might estimate less conditions since it might very well happen that no household is present both in period t and $t+k$. Formally, we estimate the available moment conditions defined as, $\mathbb{E}\left[\lambda_{i,t,k}\left(\hat{m}_{tk} - m_{tk}(\theta)\right)\right]$ where $\lambda_{i,t,k}$ equals 1 only if i is present at t and $t+k$, and is 0 otherwise, and where $\hat{m}_{t,k} = \frac{1}{I_{t,k}} \sum_{i=1}^{I_{t,k}} y_{i,t} y_{i,t+k}$, with $I_{t,k} = \sum_i \lambda_{i,t,k}$.

[46] Altonji and Segal (1996) show that the optimal weighting matrix introduces significant small sample bias. They study the small sample properties of the GMM estimator with several alternative weighting matrices and recommend using the identity matrix.

main text.

9 Proofs

Lemma 1. *Under quadratic utility, there exists no ex-ante conflict between efficiency and equity. If* $L(\mathbf{C})$ *is a network component, the ex ante Pareto-efficient risk-sharing arrangement among agents in* $\mathbf{C}$ *minimizes expected cross-sectional variability in consumption. Formally,*

$$max \sum_{i \in \mathbf{C}} \mathbb{E} u\left(c_i\right) = min \mathbb{E} \sum_{i \in \mathbf{C}} \left(c_i - \bar{c}\right)^2$$

and corresponds to solving the following mean and variance relation,

$$min \sum_{i \in \mathbf{C}} var\left(c_i\right) + \mathbb{E}\left(c_i\right)^2 \quad subject\ to \quad \sum_{i \in \mathbf{C}} c_i(\omega) = \sum_{i \in \mathbf{C}} y_i(\omega) \quad for\ every\ state\ \omega \tag{21}$$

Proof of Lemma Lemma 1

Consider the minimization of expected cross-sectional variability in consumption defined as $\mathbb{E}\left[\sum_i \left(c_i - \bar{c}\right)^2\right]$, where $\bar{c} = \frac{1}{N}\sum_i c_i$ represents the average consumption. This is equivalent to minimizing $\sum_i c_i^2 - \frac{1}{n}\left(\sum_i c_i\right)^2$. Since we have that $\sum c_i = \sum y_i$ the second term drops out of the optimization problem. As a result, the problem reduces to minimizing $\mathbb{E}\sum c_i^2$. Notice that the welfare problem under quadratic utility corresponds to minimizing $\mathbb{E}\sum_i c_i - \frac{1}{2}\gamma c_i^2$. Distributing the sum and imposing the feasibility condition that $\sum c_i = \sum y_i$, implies that the planner essentially minimizes $\mathbb{E}\sum c_i^2$. For the second statement, notice that $\sum_i \mathbb{E} c_i^2 = \sum_i var\left(c_i\right) + \mathbb{E}\left(c_i\right)^2$ follows from the definition of variance and the linearity of the expectations operator.

Proof of Proposition 1

Proof. Recall the optimality conditions given in equation (7), that characterize the dependency of shares across the network,

$$\alpha_{ij} = g_{ij}(\Lambda_j - \Psi \sum_k g_{ik}\alpha_{ik}) \tag{22}$$

where $\Psi = \frac{\mu^2+\rho}{\sigma^2-\rho}$, $\Lambda_j = \frac{\lambda_j}{2(\sigma^2-\rho)}$ and λ_j is the constant to the constraint on $j's$ outgoing shares. We will rewrite these in matrix form by defining first the vector of i's incoming shares $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})'$ and the vector of constraint multipliers $\mathbf{\Lambda} = (\Lambda_1, \Lambda_2, \ldots, \Lambda_n)'$. Let $\mathbf{G}_i$ represent the i^{th} row of $\mathbf{G}$ and let $\hat{\Psi} = -\Psi$. Then, equation (22) can be written as,

$$\left(\mathbf{I} - \hat{\Psi}\mathbf{G}_i'\mathbf{G}_i\right)\alpha_i = diag\left(\mathbf{G}_i'\mathbf{G}_i\right)\mathbf{\Lambda}$$

and, because the matrix on the left-hand-side is full-rank, we can offer the following formulation,

$$\alpha_i = \left(\mathbf{I} - \hat{\Psi}\mathbf{P}_i\right)^{-1} diag\left(\mathbf{P}_i\right)\mathbf{\Lambda} \tag{23}$$

where I set $\mathbf{P}_i = \mathbf{G}_i'\mathbf{G_i}$ for ease of notation. Now, if the value of $\hat{\Psi}$ is such that $\hat{\Psi} < \frac{1}{\nu_{max}}$ where ν_{max} is the leading eigenvalue of $\mathbf{P}_i$ then we can write the following relation,

$$\left(\mathbf{I} - \hat{\Psi}\mathbf{P}_i\right)^{-1} = \sum_{k=0}^{\infty}\hat{\Psi}^k\mathbf{P}_i^k$$

The condition on $\hat{\Psi}$ holds for all matrices $\mathbf{P}_i$ since, by the Perron-Frobenius theorem, $0 < \underset{j}{min}\sum_k \mathbf{P}_{i;kj} \leq \nu_{max} \leq \underset{j}{max}\sum_k \mathbf{P}_{i;kj}$ and $\hat{\Psi} < 0$. Now, because the matrix $\mathbf{P}_i$ is idempotent up to a scalar corresponding to the degree of i — i.e. $\mathbf{P}_i^k = d_i^{k-1}\mathbf{P}_i$ — then we can simplify the above series in the following way,

$$\sum_{k=0}^{\infty}\hat{\Psi}^k\mathbf{P}_i^k = \mathbf{I} - \frac{\Psi}{1+\Psi d_i}\mathbf{P}_i$$

Finally, it can be easily checked that $\mathbf{P}_i \cdot diag\left(\mathbf{P_i}\right) = \mathbf{P}_i$, which means we can rewrite equation (23) as,

$$\alpha_i = \left(diag\left(\mathbf{P}_i\right) - \frac{\Psi}{1+\Psi d_i}\mathbf{P}_i\right)\mathbf{\Lambda} \tag{24}$$

where we still have to solve for $\mathbf{\Lambda}$ to obtain a closed form solution of α_i. To do this notice that 24 allows us to rewrite $j's$ constraint as,

$$1 = \sum_i \alpha_{ij} = d_j\Lambda_j - \sum_i g_{ij}\left(\frac{\Psi}{1+\Psi d_i}\sum_k g_{ik}\Lambda_k\right)$$

which implies that

$$\Lambda_j = \frac{1}{d_j} + \frac{1}{d_j}\sum_{i,k} g_{ji}g_{ik}\frac{\Psi}{1+\Psi d_i}\Lambda_k$$

□

Proof of Proposition 2

Proof. If we impose the feasibility constraints on the vector equation (24) we obtain that,

$$\mathbf{1} = \sum_i\left(diag\left(\mathbf{P}_i\right) - \frac{\Psi}{1+\Psi d_i}\mathbf{P}_i\right)\mathbf{\Lambda}$$

where $\mathbf{1}$ is an $n-vector$ of ones. The properties of $\mathbf{P}_i$ means we can rewrite equation (24) as a function of the original matrix $\mathbf{G}$ in the following way,

$$\alpha_i = \left(diag\left(\mathbf{P}_i\right) - \frac{\Psi}{1+\Psi d_i}\mathbf{P}_i \right) \left(\mathbf{D} - \mathbf{G\Psi G}\right)^{-1} \mathbf{1}$$

where $\mathbf{\Psi}$ is a diagonal matrix with $\mathbf{\Psi}_{ii} = \frac{\Psi}{1+\Psi d_i}$ and $\Psi_{ij} = 0$ for all $i \neq j$. This provides a closed form solution of the constrained-efficient flows on any given network.

Finally, to arrive at the result we solve for the inverse matrix above as a series of powers of $\mathbf{G}$. The following formulation allows us to do so

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

the middle term being inverted can be expressed as a geometric series as long as ?? . Provided this is so, the following relation holds,

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

which can be understood as accumulating weighted even powers of the adjacency matrix as follows,

$$\mathbf{D}^{-1}\left(\mathbf{I} + \left(\mathbf{A} + \mathbf{A}\mathbf{D}^{-1}\mathbf{A} + \mathbf{A}\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-\mathbf{1}}\mathbf{A} + \mathbf{A}\mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-\mathbf{1}}\mathbf{A}\mathbf{D}^{-1}\mathbf{A} + \ldots\right)\mathbf{D}^{-1}\right)$$

where we set $\mathbf{A} = \mathbf{G\Psi G}$ to ease notation. So the sharing rule weights all even powers between i and j through the matrix $\hat{\mathbf{\Psi}}$ that appears between the product of $\mathbf{G}$. This can be surmised in the above expression and can be written as follows. Consider the set of all paths of length q between i and j under $\mathbf{G}$ as

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

for every $\pi_{ij} \in \Pi_{ij}^{q}\left(\mathbf{G}\right)$ let $W\left(\pi\right)$ define the weights associated to this path. It is not difficult to see that,

$$W\left(\pi_{ij}\right) = \frac{1}{d_i}\frac{\mu^2+\rho}{\sigma^2-\rho+\left(\mu^2+\rho\right)d_{i_1}}\frac{1}{d_{i_2}}\frac{\mu^2+\rho}{\sigma^2-\rho+\left(\mu^2+\rho\right)d_{i_3}}\cdots\frac{1}{d_j}$$

Finally, let M_i represent the i^{th} element of the vector $\left(\mathbf{D} - \mathbf{G\Psi G}\right)^{-1}\mathbf{1}$. Then, $M_i = \frac{1}{d_1} + \sum_j M_{ij}$, where

$$M_{ij} = \sum_{q=1}^{\infty} \sum_{\pi \in \Pi_{ij}^{2q}} W\left(\pi_{ij}\right)$$

□

Proof of Proposition 3

Proof. Assume on the contrary that $\alpha_{ij} \neq \frac{1}{d}$. A regular network has the property that $M_i\left(\Psi, \mathbf{G}_{reg}\right) = M_j\left(\Psi, \mathbf{G}_{reg}\right) = M\left(\Psi, \mathbf{G}_{reg}\right)$ for all i and j. In that case, we can write the assumption that equal sharing is not the solutions as,

$$\alpha_{ij}^{\star} = M\left(\Psi, \mathbf{G}_{reg}\right)\left(1 - \frac{\Psi d}{1+\Psi d}\right) \neq \frac{1}{d}$$

which implies that

$$M\left(\Psi, \mathbf{G}_{reg}\right) \neq \frac{1}{d} + \Psi$$

Using the result from Proposition X, we can also write $M\left(\Psi, \mathbf{G}_{reg}\right)$ as,

$$M\left(\Psi, \mathbf{G}_{reg}\right) = \frac{1}{d} + \sum_{q\in\mathbb{N}} \sum_{j\in N} \sum_{\pi_{ij}\in\Pi_{ij}^{2q}} W\left(\pi_{ij}\right)$$

where $W(\pi)$ corresponds to a particular weighting scheme for paths between i and j. So if equal sharing is not the solution for a regular network, then $\Psi \neq \sum\limits_{q\in\mathbb{N}} \sum\limits_{j\in N} \sum\limits_{\pi_{ij}\in\Pi_{ij}^{2q}} W\left(\pi_{ij}\right)$. I show next that in fact they are equal.

By the symmetry of the complete network, we know that, $W\left(\pi_{ij}^{2q}\right) = \left(\frac{1}{d}\right)^{q+1}\left(\frac{\Psi}{1+\Psi d}\right)^{q}$ for any path of length $2q$ between i and j. Finally it is just a matter of finding how many such paths there are. Let $\Pi_j^q = \underset{i}{\cup}\Pi_{ij}^q$ and Π_{ij}^q is the set of all paths of length q between i and j. Define $|A|$ as the cardinality of set A. Then we can write that,

$$\sum_{q\in\mathbb{N}} \sum_{j\in N} \sum_{\pi_{ij}\in\Pi_{ij}^{2q}} W\left(\pi_{ij}\right) = \frac{1}{d}\sum_{q\in\mathbb{N}} \left|\Pi_j^{2q}\right| \left(\frac{1}{d}\right)^q \left(\frac{\Psi}{1+\Psi d}\right)^q$$

The value of $\left|\Pi_j^{2q}\right|$ corresponds to the number of paths of length $2q$ starting from j. All paths of length $2q$ contain $2q+1$ nodes, so this is equivalent to the number of ways to assign d values to each of the $2q$ remaining values (once we fix j). This is a standard assignment problem in combinatorics and the solution is well known and equal to d^{2q}. This means that we can write the following,

$$\frac{1}{d}\sum_{q=1}^{\infty} d^{2q}\left(\frac{1}{d}\right)^q \left(\frac{\Psi}{1+\Psi d}\right)^q = \frac{1}{d}\sum_{q=1}^{\infty}\left(\frac{d\Psi}{1+\Psi d}\right)^q = \Psi$$

where the second equality comes from the convergence of the geometric series. This contradicts the original assumption that $\alpha_{ij} \neq \frac{1}{d}$ and proves the result. $\square$

Let us define a household's neighborhood centrality as a weighted average of all of its neighboring centralities as follows,

$$M_{N_i}(\mathbf{G},\Psi)=\frac{\Psi}{1+\Psi d_i}\sum_k g_{ik}M_k(\mathbf{G},\Psi) \tag{25}$$

This term appears in the constrained-efficient solution to all of $i's$ incoming shares and weights the total position of all of $i's$ neighbors by the connectivity of i. The following two lemmas derive properties of this neighborhood centrality and are used in a couple of proofs in the paper.

Lemma 2. *The WEPC of agent i can be expressed as a function of the neighborhood centralities of all its neighbors. In other words,*

$$M_i(\mathbf{G},\Psi)=\frac{1}{d_i}\left(1+\sum_k g_{ik}M_{N_k}(\mathbf{G},\Psi)\right)$$

where $M_{N_k}=\frac{\Psi}{1+\Psi d_i}\sum_k g_{ik}M_k$

Proof. Recall the two-step recursive expression of WEPC from equation 10,

$$M_i(\Psi,\mathbf{G})=\frac{1}{d_i}\left(1+\sum_{l,k}g_{ik}g_{kl}\frac{\Psi}{1+\Psi d_k}M_l(\Psi,\mathbf{G})\right) \tag{26}$$

the second term in brackets above can be rewritten as

$$\sum_j g_{ij}\frac{\Psi}{1+\Psi d_j}\sum_k g_{jk}M_k(\mathbf{G},\Psi)$$

and using the definition of M_{N_j} in equation 25 we obtain the expression. □

Lemma 3. *let $n(C)$ equal the total number of households in any connected component C, then the average neighborhood centrality over that component is always equal to Ψ. Formally,*

$$\sum_{i\in C}M_{N_i}=\Psi n(C)$$

for all C.

Proof. Using the budget constraint and our constrained-efficient solution, $\alpha^\star_{ij}$ in equation (9), we have that

$$n(C)=\sum_{i\in C}\sum_j g_{ij}\alpha_{ij}=\sum_{i\in C}\sum_j g_{ij}\left(M_j-M_{N_i}\right)=\sum_{i\in C}\sum_j g_{ij}M_j-\sum_{i\in C}M_{N_i}\sum_j g_{ij}$$

Now using the definition of M_{N_i} in equation 25, we have the following relationship

$$n(C)=\sum_{i\in C}M_{N_i}\left(\frac{1+\Psi d_i}{\Psi}-d_i\right)$$

rearranging we get the result. □

Lemmas 2 and 3 together imply the following useful result,

Lemma 4. *The WEPC of a household i that is connected to all other households in a component $C_i(\mathbf{G})$ is constant across all networks and equal to*

$$M_i(\mathbf{G}, \Psi) = \frac{1}{d_i} + \Psi$$

for all $\mathbf{G}$ whenever $d_i = |C_i(\mathbf{G})|$.

Proof. Straightforward. $\square$

Proof of Proposition 4

Proof. let h represent the center (or hub) of the star and s represent the peripheral households (or spokes). Then, define the transfer from h to s as $\alpha_{sh} = M_h(\mathbf{G}, \Psi) - M_{N_s}(\mathbf{G}, \Psi)$ using the constrained-efficient solution of equation 9 and the definition of neighborhood centrality in equation (25). By Lemma 2 we can rewrite this as

$$\begin{aligned} \alpha_{sh} &= \tfrac{1}{d_h}\left(1 + (n-1)M_{N_s} + M_{N_h}\right) - M_{N_s} \\ &= \tfrac{1}{d_h}\left(1 + M_{N_h}\right) + M_{N_s}\left(\tfrac{n-1}{d_h} - 1\right) \end{aligned}$$

and by lemma 3 we have that

$$M_{N_s} = \frac{n\Psi - M_{N_h}}{n-1}$$

which allows us to express α_{sh} only as a function of M_{N_h}.

Finally, since h by definition is connected to all other players in the network, we can use corollary 1 together with the constraint on the shares sent by h to obtain the following useful relationship between M_{N_h} and α_{sh}

$$1 - (n-1)\alpha_{sh} = \frac{1}{d_h} + \Psi - M_{N_h}$$

this implies that we can express α_{sh} uniquely as a function of parameters, Ψ, n and d_h, as follows,

$$\alpha_{sh} = \frac{1}{d_h}\left(\frac{1}{d_h} + \Psi + (n-1)\alpha_{sh}\right) + \frac{n\Psi - \frac{1}{d_h} + \Psi + (n-1)\alpha_{sh} - 1}{n-1}\left(\frac{n-1}{d_h} - 1\right)$$

rearranging we get that

$$\alpha^{\star}_{sh} = \frac{1}{1 + \left(\frac{n}{2} + 1\right)\Psi}\left(\frac{1}{d_h} + \Psi\right)$$

this proves the result for transfers from h to s. Similar steps show that Proposition 4 also holds for transfers from s to h $\square$

Proof of Proposition 6

Proof. We write down the variability in consumption as $var\left(c_i\right) \propto \alpha_i' \alpha_i$ where

$$\alpha_i = \left(diag\left(\mathbf{G}_i\right) - \frac{\Psi}{1+\Psi d_i}\mathbf{G}_i'\mathbf{G}_i\right)\left(\mathbf{D} - \mathbf{G}\boldsymbol{\Psi}\mathbf{G}\right)^{-1}\mathbf{1}$$

let $\mathbf{H}_i = \left(diag\left(\mathbf{G}_i\right) - \frac{\Psi}{1+\Psi d_i}\mathbf{G}_i'\mathbf{G}_i\right)$. Then we have that,

$$var\left(c_i\right) = \left(\left(\mathbf{D} - \mathbf{G}\boldsymbol{\Psi}\mathbf{G}\right)^{-1}\mathbf{1}\right)' \mathbf{H}_i'\mathbf{H}_i\left(\mathbf{D} - \mathbf{G}\boldsymbol{\Psi}\mathbf{G}\right)^{-1}\mathbf{1} \tag{27}$$

$\mathbf{H}_i$ is symmetric so $\mathbf{H}_i'\mathbf{H}_i = \mathbf{H}_i^2$. Equation (27) is a quadratic form on $\mathbf{H}_i^2$. The result follows from standard properties of quadratic forms. □