The Formation of Risk Sharing Networks*

Marcel Fafchamps — University of Oxford †

Flore Gubert — IRD-Paris, DIAL‡

April 2006

Abstract

This paper examines the formation of risk sharing networks in the rural Philippines. We find that geographic proximity – possibly correlated with kinship – is a major determinant of mutual insurance links among villagers. Age and wealth differences also play an important role. In contrast, income correlation and differences in occupation are not determinants of network links. Reported network links have a strong effect on subsequent gifts and loans. Gifts between network partners are found to respond to shocks and to differences in health status. From this we conclude that intra-village mutual insurance links are largely determined by social and geographical proximity and are only weakly the result of purposeful diversification of income risk. The paper also makes a methodological contribution to the estimation of dyadic models.

JEL: O1, I3

Keywords: risk-sharing; networks; dyadic models

---

*We are grateful for the excellent comments we received from Christopher Barrett, the Journal editor, two anonymous referees, and participants to the Pew Workshop on Social Dynamics and the Microeconomics of Poverty held in Bellagio in April 2005. We thank Susan Lund for making the data available. The support of the Economic and Social Research Countil (UK) and of the Pew Foundation are gratefully acknowledged. The work is part of the programme of the ESRC Global Poverty Research Group.

†Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ. Email: `marcel.fafchamps@economics.ox.ac.uk`. Fax: +44(0)1865-281447. Tel: +44(0)1865-281446.

‡DIAL, 4 rue d'Enghien, 75010 Paris. Tel : +33(1) 53 24 14 69. http://www.dial.prd.fr

1. Introduction

Much theoretical work has been done on networks by sociologists who have started thinking about networks as early as the 1960's (Mitchell 1969) and modeling them using graph theory (e.g. Raub & Weesie 1990, Weesie & Raub 2000). More recently, networks have begun receiving attention from economic theorists. Bala & Goyal (2000) and Goyal, van der Leij & Moraga-Gonzalez (2004), for instance, have studied the relationship between network architecture and underlying incentives. Kranton & Minehart (2001) have examined the restrictions on exchange that network relationships place on exchange. Genicot & Ray (2003) and Bloch, Genicot & Ray (2004) investigate the conditions under which specific network architectures are stable with respect to individual and group deviations. Recent progress has also been made – primarily by epidemiologists or under their impetus – in the modeling of large networks (Vega-Redondo 2004).

Development economists have long suspected that interpersonal relationships help shape economic exchange and agrarian institutions (e.g. Basu 1986, Bardhan 1984). This is probably because formal institutions often are weak and must be supplemented by interpersonal trust (Fafchamps 2005). This appears particularly true for risk sharing which, in addition to self-insurance via precautionary saving, has been shown to be a fundamental risk coping mechanism for the rural poor (e.g. Rosenzweig & Wolpin 1988, Townsend 1994, Ligon, Thomas & Worrall 2001, Ligon, Thomas & Worrall 2000, Fafchamps 2003). The pooling of idiosyncratic risk remains primarily informal in much of the developing world (e.g. Fafchamps 1992, Coate & Ravallion 1993, Foster & Rosenzweig 2001). In addition to risk sharing within households (e.g. Rosenzweig & Stark 1989, Dercon & Krishnan 2000), transfers and inter-personal loans constitute primary channels of risk pooling (Udry 1994). Transfers and interpersonal loans have been shown to travel primarily along long-lasting interpersonal networks (e.g. Ellsworth 1989, Lucas & Stark 1985). The same is true of labor exchange arrangements (Krishnan & Sciubba 2004).

In this paper we examine whether risk sharing networks are formed so as to maximize the potential for income risk sharing. The benefit from sharing income risk is largest when households have uncorrelated – or negatively correlated – incomes. Presumably this arises when households have very different income profiles – e.g., different occupations – and are subjected to different sources of risk – e.g., live far apart. Gains from risk sharing are thus expected to increase with social and geographical distance. However, distance also raises the cost of establishing and maintaining interpersonal links. The effect of distance on link formation is therefore a priori indeterminate.

We investigate this issue empirically using survey data collected in rural Philippines for the purpose of studying risk sharing. Fafchamps & Lund (2003) have shown that informal gifts and loans serve a risk sharing purpose but also that the extent of risk sharing is limited by the extent of interpersonal networks. Here we examine the factors determining the formation of risk sharing networks and the extent to which these networks de facto shape subsequent gifts and loans. We show that geographic proximity – possibly correlated with kinship ties – is a major determinant of interpersonal networks. This may be because it facilitates monitoring and enforcement and makes it easier to assist in case of health shock. Age and wealth differences also play an important role in the formation of risk sharing links. In contrast, occupation and income correlation are not major determinants of link formation. Reported network links have a strong effect on subsequent gifts and loans. They also affect the way gifts and loans respond to shocks, but the evidence is not statistically very strong. From this we conclude that gifts and informal loans are embedded in interpersonal relationships. These relationships are largely determined by proximity and are only weakly the result of purposeful diversification of income risk.

This paper also makes a methodological contribution to the estimation of network regressions.

We clarify the identification issues raised by dyadic regressions – that is, regressions in which each observation expresses a relationship between pairs of nodes. We also extend the concept of robust standard errors to dyadic regressions, thereby providing an easy alternative to network inference methods based on permutations or generalized least squares.

The paper is organized as follows. We begin in Section 2 by presenting our regression model and discussing various econometric issues raised by the estimation of dyadic regressions. In Section 3 we present the data and its main characteristics. Econometric results are discussed in detail in Section 4.

2. Econometric issues

We are interested in estimating a regression of the form:

$$\begin{aligned} L_{ij} &= 1 \text{ if } B(d_{ij},1) - B(d_{ij},0) - C(d_{ij}) + e_{ij} > 0 \\ &= 0 \text{ otherwise} \end{aligned} \tag{2.1}$$

where L_{ij} denotes the existence of a link between individuals i and j and d_{ij} represents the distance between i and j. The benefit from the link is denoted $B(d_{ij}, L_{ij} = 1) - B(d_{ij}, L_{ij} = 0)$, the cost of maintaining a link is denoted $C(d_{ij})$, and e_{ij} is a residual effect. The benefit from the link includes – but is not limited to – mutual insurance. A link is established if the benefit from the link exceeds the cost of maintaining it.

We interpret distance as including multiple dimensions such as spatial distance, family relatedness, shared activities, age and gender. Because of moral hazard, information asymmetries, and the ability to inflict social sanctions, we expect $C(d_{ij})$ to increase with social and geographical distance. The potential for mutual insurance is also likely to increase with d_{ij} if shocks become less correlated the more different individuals are. There is therefore a trade-off between the scope for insurance and the ability to cross-insure. As a result, it is empirically unclear whether individuals are able to establish mutual insurance links with people who are in the best position to provide such insurance. Estimating equation (2.1) is the main objective of this paper.

Equation (2.1) is a dyadic regression model. Dyadic regressions are defined as having a canonical form:

$$Y_{ij} = \alpha + \beta X_{ij} + u_{ij} \tag{2.2}$$

where i and j are individuals, Y_{ij} is an $N \times (N-1)$ matrix, and X_{ij} is a series of $N \times (N-1)$ matrices.[1] Network analysis naturally leads to regression models of this form. The estimation of dyadic regressions such as (2.2) raises two types of difficulties: identification, and inference. The first problem relates to the form in which regressors X_{ij} enter the regression. The second relates to the estimation of standard errors. We discuss these in turn.

2.1. Identification

Dyadic data contains two types of information: attributes w_{ij} of the link between i and j, such as the geographical distance between them, and attributes z_i and z_j of the nodes i and j. Regressors must enter a dyadic regression in a symmetric fashion so that the effect of (z_i, z_j) on Y_{ij} is the same as the effect of (z_j, z_i) on Y_{ji}. Dyadic regressions must therefore be written in a way that preserves this symmetry. How this is accomplished depends on whether the dyadic relationship is directional or not. Identification also depends on whether each individual i has the same number of links n_i – or degree. We discuss these two issues in turn.

A dyadic relationship is undirectional if $Y_{ji} = Y_{ij}$ for all i, j. In this case, symmetry requires

[1] The total number of possible ij pairs is N^2, but we drop the N ii pairs on the diagonal.

that regressors satisfy $\beta X_{ij} = \beta X_{ji}$. One easy way of satisfying this requirement is to specify the regression as:

$$Y_{ij} = \alpha + \beta_1 |z_i - z_j| + \beta_2 (z_i + z_j) + \gamma |w_{ij}| + u_{ij} \tag{2.3}$$

where z_i and z_j are characteristics of individual i and j thought to influence the likelihood of a link Y_{ij} between them. In the case of continuous regressors, the (2.3) is intuitive: β_1 measures the effect of differences in attributes on Y_{ij} while β_2 captures the effect of the combined level of z_i and z_j on Y_{ij}. The same formalism can be applied to the case where z_i is a dummy variable.[2]

If a dyadic relationship is directional, Y_{ij} need not equal Y_{ji}. In this case, an easy way to satisfy the symmetry requirement is to specify the model as:

$$Y_{ij} = \alpha + \beta_1 (z_i - z_j) + \beta_2 (z_i + z_j) + \gamma w_{ij} + u_{ij} \tag{2.4}$$

If z_i is a dummy variable, the (2.4) formalism still yields the desired outcome.[3]

Identification depends on degree distribution. If all individuals have the same degree, we cannot identify β_2. This follows from the fact that dyadic observations are not independent. Consequently the joint likelihood of the sample does not decompose into a product of single observation likelihoods. When all individuals have the same degree, the structure of the joint likelihood is such that only the effect of differences between observations can be identified. Showing this formally is beyond the scope of this paper but to see this intuitively, imagine we

[2] To see that this formalism also applies to the case where z_i is a dummy variable, note that in the undirectional case there are three possible configurations of (z_i, z_j): $(0,0), (1,1)$, and $\{(1,0) \text{ or } (0,1)\}$. One possible approach is to create one dummy for the $(1,1)$ configuration and another for the $\{(1,0) \text{ or } (0,1)\}$ configuration. An alternative and equivalent approach is to apply the same transformation as for continous variables in (2.3): $|z_i - z_j|$ takes two values, 0 if $z_i = z_j$ and 1 otherwise, while $z_i + z_j$ takes three possible values – $0, 1$, and 2. The effect of a $(1,1)$ configuration is given by $\beta_2/2$ gives while $\beta_1 + \beta_2$ gives the effect of a $\{(1,0) \text{ or } (0,1)\}$ configuration. The advantage of applying the (2.3) formalism to all regressors will only become apparent when we introduce degree.

[3] There are now four possible configurations: $(z_i, z_j) = (0,0), (1,1), (1,0)$ or $(0,1)$. At first glance it looks as if we could introduce three separate dummy variables. Doing so would violate the symmetry requirement, however. Indeed we need that the effect of $(0,1)$ be the negative of the effect of $(1,0)$. The (2.4) formalism ensures this since $z_i - z_j$ can take three values: $-1, 0$, and 1.

have data on monogamous couples and that z_i denotes education. By design, all individuals are paired with one and only one other individual, irrespective of their education level. We can ask the data whether educated people marry each other, but not whether educated people are more likely to be married. This means that we can identify whether differences in attributes $z_i - z_j$ affect the likelihood of a link, but not whether better educated people have on average more links. It follows that the effect of $z_i + z_j$ cannot be estimated: we can identify β_1 but not β_2. Identification of β_2 requires that individuals have different degrees, as would be the case, for instance, if the data included unmarried individuals or polygamous couples. Only then could we ask the data whether educated people are more likely to be married. Degree variation is necessary to identify level effects β_2.

2.2. Standard errors

Dyadic observations are not independent. This is due to the presence of individual-specific factors common to all observations involving that individual. It is thus reasonable to assume that $E[u_{ij}, u_{ik}] \neq 0$ for all k and $E[u_{ij}, u_{kj}] \neq 0$ for all k. By the same reasoning, we also have $E[u_{ij}, u_{jk}] \neq 0$ and $E[u_{ij}, u_{ki}] \neq 0$.[4] Provided that regressors are exogenous, applying OLS to (2.3) and (2.4) yields consistent coefficient estimates but standard errors are inconsistent, leading to incorrect inference.

Robust standard errors must correct for cross-observation correlation in the error terms involving similar individuals. To obtain such robust standard errors, we extend the method that Conley (1999) developed to deal with spatial correlation of errors. Conley's method is itself an extension of the robust covariance matrix popularized by White and applied to time series by

[4] This situation bears some formal resemblance to random effects models with two-way error components discussed for instance by Baltagi (1995), except that here we have four-way random effects.

Newey and West. The formula for network corrected covariance matrix is of the form:

$$AVar(\widehat{\beta}) = \frac{1}{N-K}(X'X)^{-1}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}\frac{m_{ijkl}}{2N}X_{ij}u_{ij}u'_{kl}X_{kl}\right)(X'X)^{-1} \tag{2.5}$$

where β denotes the vector of coefficients, N is the number of dyadic observations, K is the number of regressors, X is the matrix of all regressors, X_{ij} is the vector of regressors for dyadic observation ij, and $m_{ijkl} = 1$ if $i = k, j = l, i = l$ or $j = k$, and 0 otherwise.[5] The only structure imposed on the covariance structure is that $E[u_{ij}, u_{ik}] \neq 0$, $E[u_{ij}, u_{kj}] \neq 0$, $E[u_{ij}, u_{jk}] \neq 0$ and $E[u_{ij}, u_{ki}] \neq 0$ for all k but that $E[u_{ij}, u_{km}] = 0$ otherwise.[6] Formula (2.5) also corrects for possible heteroskedasticity.

Monte Carlo simulations indicate that standard errors corrected for dyadic correlation can be much larger than uncorrected ones. The bias is particularly large when the average degree is high. Correcting standard errors is thus essential when estimating any dyadic regression. In our case, the magnitude of the correction is relatively small because the average degree is low.

Other methods have been devised to conduct inference on network data. One such method relies on permutation methods popularized by Good (2000). This method was first applied to network analysis by Hubert & Schultz (1976) and subsequently refined by Krackhardt (1987) and Nyblom, Borgatti, Roslakka & Salo (2003). Instead of correcting standard errors, permutation methods correct p-values directly. This procedure is known as Quadratic Assignment Procedure or QAP in the literature (Hubert & Schultz 1976). This approach has gained much popularity among sociologists who typically compute QAP p-values using a linear probability model. We believe our method to be statistically more efficient since it does not rely on bootstrapping. The

[5] By construction, all observations where $j = i$ or $k = l$ are identically zero and hence are omitted. Division of the inner term by 2 corrects for the double counting implied by the simple way we have written the formula.

[6] Computation of (2.5) as written is very computer-intensive. It is however possible to take advantage of specific Stata commands and of the structure of the m_{ijkl}'s to reduce computation to a small number of matrix manipulations. To apply formula (2.5) to logit, $(X'X)^{-1}$ needs to be replaced by an expression that depends on the scores.

first version of this paper used QAP to derive p-values. Inference was similar to using dyadic standard errors.[7]

3. The data

Having presented the conceptual framework and discussed econometric issues, we now describe the data. Two types of sampling approaches have been used to collect network data. In the two-step approach, the researcher samples a population at random, obtains information about their network links, and surveys these links in a second step. This approach is particularly interesting to study in-degree and network centrality – for instance, to identify people used as source of finance by many. The problem with this approach is that the second step nodes are not selected randomly. This complicates inference.

In the one-step approach, a random sample of the population is selected and information is collected on the links among them. Depending on time and resources, one may collect information on all existing links, on the most important ones, or a random selection of them, as Conley & Udry (2001) have done for instance. This approach offers the advantage that both nodes in any link are randomly selected. This simplifies the econometric estimation of dyadic regressions. Since in-degree and centrality are not the focus of our analysis, we use the one-step approach.

A survey was conducted in four villages in the Cordillera mountains of northern Philippines between July, 1994 and March, 1995 (Lund 1996). A random sample of 206 rural household was drawn after taking a census of all households in selected rural districts. These households are dispersed over a wide area; most can only be reached by foot. Three interviews were conducted with each household at three month intervals between July 1994, just after the annual rice

[7] A GLS estimator for dyadic data – called P2 – has also been developed by sociologists (Dekker, Krackhardt & Snijders 2003). This method is sensitive to heteroskedasticity, which is why we prefer to keep OLS and logit estimates and correct standard errors directly (Wooldridge 2002).

harvest, and March 1995, after the new rice crop had been transplanted.

Sample households derive most of their income from non-farm activities (Fafchamps & Lund 2003). There are many skilled artisans in this area, and their wood carvings, woven blankets, and rattan baskets supply a growing tourist and export trade. Unearned income – mostly land rentals – is not negligible but very unevenly distributed across households, as is often the case with asset income. Although nearly all households operate their own farm, the majority do not produce enough grain to meet annual consumption needs. Sales of crops and livestock account for a minute fraction of total income.

The data indicate that differences in income per capita across households are significantly correlated with differences in wealth ($\rho = 0.16$; p-value$= 0.000$) and education levels ($\rho = 0.19$; p-value$= 0.000$). They are also negatively correlated with differences in distance from the road. This means that individuals located close to each other tend, on average, to have less similar incomes. The effect is quite small, however ($\rho = -0.05$). We also find that households with different levels of education are less likely to be engaged in the same occupation.

At the beginning of the survey, each household was asked to identify a maximum of four individuals on which it could rely in case of need or to whom the respondent gives help when called upon to do so.[8] These individuals constitute what we call the network of insurance partners of each household. Approximately 939 network members are identified by the survey. Of these, 189 or 20.1% are (members of) households already in the survey. In 68 of these cases, both respondents cite each other as network partners. In the rest of the cases, only one respondent cited the other household as part of their network. This is not too surprising given the question that respondents were asked to answer: that A matters to B does not necessarily

[8] In practice, respondents listed on average 4.6 individuals, with a minimum of 1 and a maximum of 8. This is because in a number of cases respondents refused to rank individuals they regarded as equivalently close to them. In such cases, enumerators were instructed to accommodate respondents rather than antagonize them.

implies that B matters to A. Still, it serves as reminder that answers to the question do not capture all the relationships that respondents are involved in. The network partners we have identified probably constitute the nucleus of a larger, more diffuse network which is difficult to quantify. The first panel of Table 1 shows that most insurance partners are close family members, e.g., children or siblings. The second panel of the table indicates that most of them (63.3%) reside in the same village (*barangay*).

In the analysis presented here, we focus on the link between sample households, which are all residents of the four studied villages. As indicated by Table 1, surveyed households also have links with individuals outside the study area, such as migrants in distant cities or traders in nearby towns. It is quite conceivable that these individuals play an important role in risk sharing, as pointed out, for instance, by Rosenzweig & Stark (1989) or Lucas & Stark (1985). Our analysis should thus be viewed as a study of network formation among villagers. This is a valid focus, given the assumption often made in the literature that villagers pool idiosyncratic risk. But it is a limitation of our approach that should be kept in mind.

Information was also collected on all debts and gifts. Respondents were asked to list all loans and transfers taking place within the last three months of each survey round. Great care was taken to collect data on all possible in-kind payments and transfers, including crops, meals, and labor services. The identity of the partner was recorded for each transaction, whether the person was identified as a 'linked' individual or not. Using these data Fafchamps & Lund (2003) have shown that gifts and loans serve in part to deal with income and health shocks and that shocks to network partners affect households' ability to borrow or receive gifts. Building on the work of Udry (1994), Fafchamps & Gubert (2002) have further shown that the repayment of loans between friends and neighbors is contingent on shocks, further confirming the insurance role of gifts and loans in the study area.

4. Empirical estimation

We now turn to the estimation of equation (2.1). We proceed in two steps. First we estimate a reduced form in which we control for geographical distance w_{ij} and various factors z_i and z_j thought to influence the level, variance, and covariance of income. Secondly we estimate a structural model of link formation in which the level, variance, and covariance of income are instrumented using the z_i's as regressors. We begin with the first set of results.

4.1. The determinants of network formation

For our reduced form analysis, we regress L_{ij} on various measures of social and geographical distance.[9] Descriptive statistics for variables used in the paper are presented in Table 2. The dependent variable L_{ij} is equal to 1 if household i cited household j as source of assistance, and 0 otherwise. Since i can cite j without j citing i, L_{ij} is directional.

Let us start by saying that we do not have information on kinship or relatedness for all household pairs in the sample. We only have that information for linked pairs. The reason is that collecting such data is extremely time-consuming since each household has to be asked about its family relationship with each of the other 205 households in the survey. We feared losing many respondents by including such a cumbersome module in the survey.[10] This means that we cannot formally investigate whether blood or marriage ties are a strong link determinant,

[9] It is conceivable that households form links with someone only to link up, through that person, with somebody else. A proper treatment of this issue goes well beyond the scope of this paper because it involves modeling complex strategic interactions between households (e.g. Genicot & Ray 2003, Bloch, Genicot & Ray 2004). Even in such an environment, however, it is possible to ask whether there are individual characteristics that help predict link formation.

Fafchamps and Lund discuss the issue of 'friction' in the risk sharing network. They point out that if transfers circulate costlessly through the network, the network architecture does not matter. Their evidence indicates that the architecture does matter, suggesting the presence of friction. In that case, linking up to someone through a proxy is less effective than linking up with the person directly. When asked to state the reason for taking a loan, some respondents state that the money is meant to help someone else, but these cases are rare.

[10] More efficient techniques for eliciting kinship information have now be devised. Recently de Weerdt (2002), for instance, managed to collect such information for all 120 households in a Tanzanian village. This issue deserves more research.

as we suspect.[11] To the extent that relatives reside near each other, as seems to be the case in the study area, geographical proximity captures some of the effects of kinship.

Geographical distance is measured by two variables. The first one is a dummy variable taking value 1 if both households i and j reside in the same *sitio*, a small cluster of 15 to 30 households. The anthropological literature describes *sitios* as traditional community groups composed mainly of kin.[12] Living in the same *sitio* is thus related to kinship. The second variable captures the difference between i's and j's distance to the nearest road, provided they reside in the same sitio. Presumably, if households in the same sitio are at the same distance from the nearest road, they are close geographically.[13] In addition to its correlation with kinship, geographical proximity is also expected to help alleviate monitoring and enforcement difficulties – and hence to lower the cost of maintaining a link. To the extent that incomes are spatially correlated, it also reduces expected gains for income pooling. But it opens more opportunities for helping each other in case of health shocks: proximity makes it easier to provide home care, to comfort the bereaved, and to assist with visits to health facilities.

We focus on six dimensions of social distance: occupation, household size, age, health, education, and wealth. We expect benefits from the pooling of income risk to be largest between people with different occupations – and especially high between farmers and non-farm workers. Farming risk is primarily determined by weather conditions and pest infestation. Non-farm

[11] It would also have been interesting to contrast male and female networks. Because nearly all respondents are male household heads, our data does not allow an investigation of this issue.

[12] E.g., http://countrystudies.us/philippines/42.htm: "In the rural Philippines, traditional values remained the rule. The family was central to a Filipino's identity, and many sitios were composed mainly of kin. Kin ties formed the basis for most friendships and supranuclear family relationships. Filipinos continued to feel a strong obligation to help their neighbors– whether in granting a small loan or providing jobs for neighborhood children, or expecting to be included in neighborhood work projects, such as rebuilding or reroofing a house and clearing new land. The recipient of the help was expected to provide tools and food. Membership in the cooperative work group sometimes continued even after a member left the neighborhood. Likewise, the recipient's siblings joined the group even if they lived outside the sitio. In this way, familial and residential ties were intermixed."

[13] This is of course a noisy measure since, in sitios located very close to the road, it is possible for two people to be far apart and yet to be the same distance to the road. But in such sitios distance to the road is short for everyone, so distance differences are close to zero. This is less of an issue for sitios located away from the road. On average surveyed households are located 17.5 minutes of walking distance from the nearest road.

income risk is largely influenced by demand for crafts by traders and tourists visiting the area. Consequently we expect both sources of risk to be uncorrelated with each other. The data indeed show a very low – and non-significant – 0.06 correlation between farm and non-farm income. If farming and non-farm incomes have a large collective risk component, pooling income within each occupation should be fairly ineffective at reducing risk. To the extent that network formation is driven by the desire to maximize gains from income pooling, links should be formed primarily between people with different occupations.

The simplest measure of occupation we use in our analysis is a dummy variable that takes value 1 if the main occupation of the household head is agriculture, and 0 otherwise. This variable captures the main occupational divide in the study area. We also experiment with a finer breakdown of occupation that combines information on primary and secondary activities of all household members. The number of working members in the household is also included as regressor because household with more income earners can be more diversified and hence less dependent on outside assistance (e.g. Binswanger & McIntire 1987, Fafchamps 2003).[14] We do not control for household size per se in order to abstract from child fostering, which has been documented to play a risk sharing role (e.g. Evans 2004, Akresh 2004).

As pointed out earlier, income pooling is not the only form that risk sharing can take. Taking care of the sick and elderly is another. Health status is captured by an index variable that takes value 1 if the household head is in good health, 2 if the head is often ill, 3 if he or she is chronically ill, and 4 if he or she is disabled. Presumably individuals with worse health need more assistance but are also less able to reciprocate. Differences in terms of age also raise the potential for risk pooling: presumably, young households with many children face different health risks from the

[14] In the analysis presented here we take the number of working adults in the household as given. Given the short-term nature of our analysis (a total duration of 9 months), this is a reasonable approximation. Over a longer time horizon, Frankenberg, Smith & Thomas (2003) have shown that households adjust their size in response to large macro shocks.

elderly. Differences in age are also likely to be associated with differences in lifestyle, perhaps reducing social interaction across groups. Again, if benefits from pooling risk across categories outweigh the cost of linking up, we expect more links between different age groups.

The remaining measures of social distance are education and wealth. Households with better education and more wealth probably have higher incomes as well. To the extent that absolute risk aversion is decreasing with income, as is customarily assumed, households with high average income are in a good position to offer income insurance to poorer households (Fafchamps 1999). Risk sharing may also have a redistribution component and the rich may be expected to help the poor, irrespective of risk sharing. For these two reasons, establishing links with richer and better educated households is attractive to poor households.[15] Rich households, in contrast, would see less need for links with poor households – or may not even see them as source of insurance.

Because insurance affects income and thus the ability to accumulate assets, wealth is potentially endogenous to the network formation process: households with better networks may accumulate more wealth. For this reason we instrument individual wealth using variables that predate the purposive formation of insurance links, namely: the education of head; the value of the inheritance of the head and spouse; whether the head was born in the village of current residence; whether the household head is male; and the number of siblings of the head and spouse. All instrumenting regression results are presented in Appendix. Since wealth enters dyadic regressions in difference and sum, we estimate two instrumenting regressions, which are presented in Table A1 together with t-values based on the dyadic standard errors obtained from

[15] Education is also as a possible source of insurance. In poor societies such as the one we study, knowledge is valuable, particularly regarding contacts with the outside world (e.g., government authorities, cooperative bank, health facilities, traders, extension agents). To rural dwellers, educated households may thus be seen as providing some protection against abuse in dealings with the outside world. Educated households may also be less vulnerable (Glewwe & Hall 1998) and recover more easily from collective shocks (Barrett, Sherlund & Adesina 2004). For this reason, we expect gains from risk sharing to be higher between households with different education levels. However, differences in education level may also increase social distance and make socialization more difficult (Mogues & Carter 2005).

(2.5). Given that they are estimated from the same data, their coefficients are very similar. We see that instruments have a strong predictive power, particularly inherited wealth. Predicted wealth variables from these regressions are used in lieu of actual wealth in all the analysis that follows.

Our first set of regression results is presented in Table 3. By construction, geographical distance variables are undirectional; as they are link attributes, they enter the regression as such. In contrast, each individual attribute is used to construct two regressors of the form $z_i - z_j$ and $z_i + z_j$. Since the dependent variable L_{ij} is directional, regressors $z_i - z_j$ enters the regression as such, not in absolute value.

As explained in the econometrics section, we can only estimate the coefficient of $z_i + z_j$ regressors if individuals have different degree. Although the number of network links varies a bit across respondents,[16] it is important to realize is that the survey did not seek to measure the number of links of each respondent. We thus do not have a strong basis for identifying level effects $z_i + z_j$. Although the degree variation present in the sample makes identification possible in practice, the resulting estimates may not be reliable. For this reason, we estimate our model with and without level effects.

As explained earlier, the data come from four villages or *barangay*. We found no link between households across village boundaries: being in different villages perfectly predicts the absence of a link. There is therefore no point in including pairs of individuals from different villages in the estimation. For this reason, we only include pairs that come from the same barangay. This explains why the reported number of observations is less than $N^2 - N = 42230$.[17]

[16] This variation is due to three sources. First, some respondents could not give four names. Second, some respondents refused to limit their response to four. Third, even for those households who listed four links, degree variation arises when we restrict the attention to network partners who are themselves in the sample.

[17] The dyadic standard error formula is adjusted to take this into account. This is accomplished by computing (2.5) for each village separately, summing over the four villages, and dividing by the total number of observations.

The first column of Table 3 presents our first set of logit estimates without level effects.[18] Robust dyadic standard errors are reported throughout. Village (*barangay*) dummies are included to control for possible village effects. Geographical effects appear strongly significant: respondents are much more likely to cite as a mutual insurance link someone residing in the same sitio. Conditional on living in the same sitio, respondents are also more likely to cite someone close to them within the sitio. Geographical proximity is unambiguously a strong predictor of network links. As we pointed out earlier, spatial proximity reduces the scope for pooling agronomic risk (pests, floods, landslides) but it facilitates monitoring and enforcement, especially given the fact that it is correlated with kinship. It also makes it easier to look after a sick neighbor and thus enhances the scope for pooling health risk.

The age difference variable is significant: younger heads of household are more likely to mention a link with an older household. This is consistent with the pooling of health risk, although it could also result from life cycle effects or intergenerational altruism. Wealth is also significant: consistent with expectations, households are more likely to mention as source of insurance households that are richer than themselves.

Contrary to expectations, education, occupation, and the number of working adults are not significant. The big surprise is that occupation is not significant: households primarily involved in farming activities are *not* more likely to be linked with non-farmers. Surveyed households do not appear to form the mutual insurance links with other villagers that would maximize the gains from pooling idiosyncratic *income* risk. Intra-village links thus appear different from inter-region links, such as links with distant migrants, which have been documented to serve an income diversification function (e.g. Lucas & Stark 1985, Rosenzweig & Stark 1989, Lauby &

[18]The reader may worry that logit may not be appropriate in this case given the very small proportion of non-zero values of the dependent variable. To investigate whether this is a cause for concern, we reestimated the model using an extreme value distribution instead of a logistic distribution. Virtually identical results obtain.

Stark 1988).

To check the robustness of our results, we add level effects and reestimate the model. As emphasized earlier, the coefficients of level effects may not be estimated reliably in our data, so we will not discuss their interpretation in much detail. Regression results are presented in the second column of Table 3. Our findings are unchanged. The only significant $z_i + z_j$ variable is the number of working adults in the household: links are less likely to be reported between households with many working age adults. This is consistent with the view that large households themselves serve to pool risk, thereby reducing the need for networking (e.g. Binswanger & Rosenzweig 1986, Binswanger & McIntire 1987, Fafchamps 2003).

Another possible source of concern is that households may locate close to other households with whom they wish to pool income risk. This could explain why spatial proximity is strongly significant while occupation and education are not. To investigate this possibility, we reestimate the model with only the households whose heads are residing in the village of their birth. We correct for self-selection using a probit selection equation shown in Table A2 in Appendix. The dependent variable is 1 if the heads of both households i and j are living in the village of their birth, 0 otherwise. Birth order is used as instrument. Education, age and gender are also included as pre-determined regressors. Given the local culture (Quisumbing 1994), we expect first-borns to remain close to their parents, and thus to live in the village of their birth. Results indicate that birth order is a significant predictor of residence in birth village. Education is also significant, with the expected sign: educated individuals are more likely to move out of rural areas, at least for a while.

Results from the selection equation are used to construct a Mills ratio for each pair of respondents i and j. This Mills ratio is then included in the dyadic regression as additional regressor. Regression results are shown in the last column of Table 3. Although the number of observations is much smaller, results are basically unchanged except that they are slightly less significant. The Mills ratio is far from significant, suggesting that self-selection is not a source of concern. These findings suggest that our results are not the consequence of endogenous household placement.

Next we investigate whether the non-significant result for occupation is due to mis-specification. To this effect, we replace the farmer dummy with a more detailed description of the activities undertaken by all members of the household. Regression results are shown in Table 4, with and without level effects. We see that identical results obtain: wealth, age, and the geographical variables remain basically unchanged in magnitude and significance while none of the occupation variables is even remotely significant.

4.2. Network formation and insurance

The analysis so far has found no evidence that surveyed households form risk sharing links primarily with other villagers with whom gains from income risk pooling would be maximized. This may be because the reduced form approach obscures the income pooling motive. To investigate this possibility, we estimate an alternative model in which we control directly for the correlation in the income of i and j. If respondents form links with people who have an income less correlated with theirs, the coefficient of income correlation in (2.1) should be negative. This is the basis for our testing strategy.

We begin by constructing a measure of income y_i^t for each household in each of the three survey rounds. The income variable includes earnings from jobs held in the last three months (e.g., casual labour, woodcarving, basket making, blanket weaving), unearned incomes received in the last three months (e.g., rents, pensions), and earnings from the sale of crops and livestock in the last three months. The latter component of income is minimal in the studied area. The imputed value of own agricultural production is not included in the computation of the income aggregate since it does not change over the duration of the survey and hence does not contribute to income correlation measured over the three rounds. It is important to keep in mind that, given data constraints, variable y_i^t does not include year-to-year variation in agricultural income. But y_i^t measures short-term fluctuations in non-farm income. Given that non-farm income is a dominant source of livelihood in the study area, y_i^t is still a useful variable to look at: households have to smooth consumption over time, and mutual insurance links with other villagers may help them do that.

Using y_i^t, we compute the correlation ρ_{ij} of income between all ij pairs. Needless to say, since ρ_{ij} is computed using only three periods, it is estimated with much measurement error. To correct for this, we instrument ρ_{ij} using all the regressors appearing in the last column of Table 4. Since ρ_{ij} is undirectional, difference regressors all appears in absolute value. Results are shown in the first column of Table A3 in Appendix. As anticipated, households with more working members have less correlated incomes. But contrary to expectations we find that incomes are *more* correlated between households with a different level of education and with a different number of members involved in farming or other self-employment activities. Other coefficients are non significant and the predictive power of the regression is quite low ($R^2 = 0.03$).

We reestimate equation (2.1) with predicted income correlation as additional regressor.[19] Results are presented in the first column on Table 4. We see that $\widehat{\rho}_{ij}$ has a positive sign, but is not significant.[20] Other results are unchanged. This finding constitutes additional evidence that links are not formed preferentially with individuals who have a less correlated income.

This result could be because we have omitted the expectation and standard deviation of

[19] To be more precise, we follow Smith & Blundell (1986) and include the residuals from the instrumenting equation as additional regressors. Although this technique has only been proven valid for probit regressions, by analogy it should also work better in logit.

[20] If we use the covariance of income instead of the correlation, results are similarly non-significant.

income from the regression. Indeed, in a general model of risk sharing, the mean and variation of income also affect the utility gain from risk sharing and hence possibly the matching process. To investigate this, we construct the mean μ_i and standard deviation σ_i of y_i^t from individual round-level data and instrument them using the same regressors. Instrumenting regressions are presented in last four columns of Table A3 in Appendix. To keep in line with the form of equation (2.1), we separately instrument the sums $\mu_i + \mu_j$ and $\sigma_i + \sigma_j$ and differences $\mu_i - \mu_j$ and $\sigma_i - \sigma_j$, but this is not necessary.[21] Results show that, as anticipated, wealth and the number of working adults are strong predictors of income and that, in the study area, wage earners have lower incomes on average. This probably reflects that nature of wage employment among villagers in the study area. As anticipated, we find that households with more income earners have a less variable total household income while households in which the head is in poor health have less variable income – probably because income is lower.

We then enter the predicted residuals from the sums and differences regressions as additional regressors in equation (2.1). Results are presented in the second column of Table 5. Income covariance remains positive and non-significant. Contrary to expectations, the likelihood of a link is seen to decrease with the (sum of) the standard deviation of income: presumably, households with more risk income would be more keen to form mutual insurance links. We also see that links are less likely between households with different levels of income variation. These findings are again inconsistent with a mutual insurance motive for link formation: if villagers form links to pool income risk, other things being equal individuals with more variable income would seek to link up with individuals whose income is less variable. We also estimated equation (2.1) using uninstrumented correlation, mean, and standard deviation of income. Income correlation results

[21] In the sum regressions, only sums of regressors enter while in the difference regressions only difference regressors enter. Results are nevertheless very similar, as expected. Virtually identical results obtain if we first instrument μ_i and σ_i, obtain the predicted values, and then take the sum and difference of the predicted values to construct the necessary regressors.

remain unchanged.

It is conceivable that we have failed to find evidence of risk sharing because we have looked at the wrong risk measure. To investigate this possibility, we broaden our definition of risk to include other types of shocks. The survey collected information on a variety of income and consumption shocks, such as crop failure, unemployment, sickness, and funerals. Moreover, respondents were asked to provide a summary assessment of their financial situation in each survey round. Responses range from -2 for very good to +2 for very bad. Because this assessment may be endogenous to network formation, it is instrumented using objective shock measures, village-time dummies, and household fixed effects as instruments. We then reestimate equation (2.1) as in Table 5, replacing income correlation with covariance in summary assessments. The regressor of interest is the (instrumented) covariance of the summary assessment variable: if its coefficient is negative, this constitutes evidence that villagers form links to share risk. Results, not shown here to save space, instead yield a significant *positive* effect of the covariance of link formation. Put differently, villagers who are linked have a higher level of covariance between their summary assessments. This is again inconsistent with the idea that links are formed to maximizes the potential for risk sharing.

4.3. The role of network links

Could it be our conclusions are erroneous because we have been looking at the wrong measure? What if reported links were irrelevant but mutual assistance did take place in a way that is consistent with the maximization of gains from risk sharing? To see how this situation could arise, imagine that strong bonds exist between most villagers. When faced with a shock, households simply go to the person most able to assist them. Asked by enumerators to provide four names, respondents may have listed the first people that came to mind, such as immediate neighbors.

But these people need not play a practical role in sharing risk because the actual network is much larger.

To show that our results are not driven by such underlying process, we need to test two hypotheses: first, that ex post risk sharing follow a pattern similar to those of reported links; and secondly that reported links are not irrelevant, that they play a role in the sharing of risk. To investigate the first hypothesis, we examine all gifts and loans received by respondent households in rounds 2 and 3 of the survey. Fafchamps & Lund (2003) have shown that, in the study area, gifts and loans play an important risk management function. We drop gifts and loans received in round 1 to avoid the spurious correlation that could arise if people interviewed in round 1 listed as network links individuals who have helped them in the recent past.

We begin by noting that a little under half of all gifts and loans come from linked individuals. This shows that linked individuals are not the only possible source of help, suggesting that the actual network is indeed larger than the reported network. But the proportion of gifts and loans coming from linked individuals is nevertheless much larger than their share of the total number of possible pairs, suggesting that reported links are not irrelevant. To investigate this formally, we estimate a model of the form:

$$G_{ij}^t = \alpha + \beta_0(z_i - z_j)L_{ij}^{t_1} + \beta_1(z_i - z_j) + \beta_2(z_i + z_j) + \gamma w_{ij} + u_{ij} \tag{4.1}$$

where G_{ij}^t denotes the value of all gifts (or loans) received by i from j in round $t = t_2$ and t_3.[22] Variables z_i and w_{ij} are as in Table 3. We also include village-time dummies to control for shocks that are common to all villagers. The interaction terms $(z_i - z_j)L_{ij}^{t_1}$ are included to test

[22] To minimize recall bias, we construct G_{ij}^t by including gifts (and loans) reported as received from j by i as well as gifts (and loans) reported as given by j to i. When i and j report something different, we keep the largest value of the two. To reduce the weight of outliers while avoiding losing all 0 observations, we take as dependent variable the log of $(1+G_{ij}^t)$.

whether reported links are irrelevant. If wealth differences are found to explain gifts received, this can be taken as evidence that gifts serve a redistribution purpose.

Results for gifts are shown in Table 6. Geographical proximity variables are strongly significant in both the gift and loan regressions, confirming earlier results. We find no evidence that gifts flow more between individuals with different occupation, ruling out the idea that respondents make a deliberate effort to share risk across occupations. Wealth differences are also non-significant, indicating that on average gifts do not serve a redistributive purpose. But we find that individuals who are less healthy – i.e., have a higher health index – are more likely to receive gifts. We also see that the existence of a reported link triggers gifts to unhealthy individuals that are 200 times larger – suggesting that these links are not irrelevant after all.

Results for loans, which are presented in the second column of Table 6, show that occupation does matter: farmers are less likely to receive loans from non-farmers – or alternatively non-farmers are more likely to receive loans from farmers. This effect is magnified 27 times if a link was reported between the two individuals. These results again confirm that reported links are not irrelevant. They also show that informal loans are affected by occupational complementarities.

We also need to show that reported links do serve a risk sharing purpose. To investigate this, we estimate a model of the form:

$$G_{ij}^t = \alpha + \theta_1(S_i^t - S_j^t) + \theta_2(S_i^t - S_j^t)L_{ij}^{t_0} + \theta_3 L_{ij}^{t_0} + \beta_1(z_i - z_j) + \gamma w_{ij} + u_{ij} \tag{4.2}$$

where G_{ij} as before denotes gifts (or loans) given by j to i, , and S_i^t and S_j^t measure shocks that individuals i and j at time $t = t_2$ and t_3. If gifts or loans serve a risk sharing purpose with villagers at large, we should observe gifts from people experiencing a good shock (low value of S) to people experiencing a bad shock (high value of S) – and hence $\theta_1 > 0$. If gifts and loans serve a risk sharing purpose only with reported link individuals, then we should observe $\theta_1 = 0$

and $\theta_2 > 0$. Because a link may favor gifts for purposes other than risk sharing, we control for $L_{ij}^{t_0}$ separately to avoid spurious results.

Our shock measure S_i^t is taken from responses to a subjective assessment question. Respondents were asked in each round to rank their situation relative other the recent past. Responses range from -2 for very good to +2 for very bad. Based on subjective assessment, this variable is of course subject to measurement error. To correct for this, we instrument it using responses to objective risk factors, such as as whether any member of the household experienced an acute or mild illness in the three months preceding the interview, whether any member of the household became unemployed, and whether the household was obligated to incur a large ritual expense (such as a funeral). The instrumenting regression is shown in Table A4 in Appendix. We see that all four objective shock variables have the anticipated sign and are strongly significant. We then use the predicted value $\widehat{S}_i^t$ from this regression in lieu of S_i^t in equation (4.2).[23]

Results are presented in Table 7 for gifts and loans. We see that θ_1 is not significantly different from 0 in both regressions but θ_2 is positive and is significant at the 10% level in the gift equation. While the result is not very strong, it nevertheless provides some support to the idea that reported links do serve a risk sharing purpose. We find that reported links are a strong predictor of gifts and loans received, irrespective of shocks. Reported links are thus not irrelevant, they do capture an important dimension of social interactions among villagers. We also estimated equation (4.2) with individual shock variables, but results are not generally not significant. This is probably because we do not have enough pairs with non-zero gifts and loans, making identification difficult. We do however find that acute sickness is nearly significant at the 10% level in the gift regression (t value of 1.64).

[23]When we use the actual value S_i^t, results are non-significant. This is consistent with the atenuation biased caused by measurement error.

5. Conclusion

In this paper we have examined the determinants of risk sharing links among households. It is indeed increasingly recognized that informal risk sharing plays a major role in the way the rural poor deal with risk (e.g. Rosenzweig & Wolpin 1988, Townsend 1994, Ligon, Thomas & Worrall 2001) and that interpersonal networks facilitate informal risk sharing (e.g. Fafchamps 1992, Dercon & de Weerdt 2002, Fafchamps & Lund 2003, Dercon & Krishnan 2000).

Theory predicts that social and geographic distance between households raises the potential benefits from risk pooling but also the cost of establishing and maintaining interpersonal links. The effect of distance on link formation is therefore theoretically indeterminate. If costs rise sufficiently rapidly with distance, the pooling of risk across households with different income profiles will not be achieved. The efficiency of informal risk pooling thus depends on the way risk sharing networks are formed.

We investigated this issue empirically using a specifically designed survey in rural Philippines. We examined which dimensions of social and geographical distance predict the existence of risk sharing relationships. We found that geographic proximity is a major determinant of interpersonal relationships, possibly because it captures kith and kin relationships and facilitates monitoring and enforcement. Age and wealth differences also play an important role in the formation of risk sharing links. In contrast, occupation is not a determinant of network links. Neither is income correlation. We also find that reported network links have a strong effect on subsequent gifts and loans. Gifts in kind and in cash between network partners are found to respond to shocks, but the evidence is not statistically very strong.

These findings suggest that surveyed households do not form links that maximize potential gains from sharing income risk. Why this is the case is unclear, but the body of the evidence presented here suggests that link maintenance costs – proxied by geographical distance – prevent households from forming links that would be optimal from the point of view of income risk sharing. However, we should stress that links with individuals outside the four surveyed villages are not included in the analysis presented here. We suspect that these links are much more important for income smoothing than intra-village links.

The strongest evidence of intra-village risk sharing relates to health risk. Households in which the head is afflicted by health problems receive more gifts from other villagers, especially from network partners, and households in which one member suffered an acute illness receive more gifts from network partners as well, albeit the effect in only marginally significant. Combined with the fact that age differences are an important determinant of village network link, this constitutes impressionistic evidence that mutual insurance networks between villagers may be formed more with health risk than income risk in mind – see Dercon & de Weerdt (2002) for a similar point. This issue deserves more research.

The literature has shown that income risk is not efficiently pooled in village economies (e.g. Townsend 1994, Ligon, Thomas & Worrall 2001, Foster & Rosenzweig 2001, Fafchamps & Lund 2003). This paper suggests that households do not appear to purposefully form links with villagers who have a different income profile. In these conditions, it is hardly surprising that efficient income risk sharing has consistently been rejected among the rural poor. Having uncovered one of the reasons why efficiency is not achieved, the challenge is now to find ways of encouraging risk pooling across occupations and income profiles.

This paper also makes a methodological contribution to the burgeoning empirical literature on economic networks (e.g. Krishnan & Sciubba 2004, Goyal, van der Leij & Moraga-Gonzalez 2004, Fafchamps, Goyal & van der Leij 2005). First we clarified identification issues in dyadic data, especially with respect to directed networks and degree distribution. Second we facilitated inference on network processes by applying the well-known concept of robust standard errors to dyadic data. These methodological improvements should assist other researchers working with dyadic data in general, and with network data in particular.

References

Akresh, Richard. 2004. "Risk, Network Quality, and Family Structure: Child Fostering Decisions in Burkina Faso." (mimeograph).

Bala, Venkatesh & Sanjeev Goyal. 2000. "A Non-Cooperative Model of Network Formation." *Econometrica* 68(5):1181–1229.

Baltagi, Badi H. 1995. *Econometric Analysis of Panel Data*. Chichester: Wiley.

Bandiera, Oriana & Imran Rasul. 2002. Social Networks and Technology Adoption in Northern Mozambique. Technical report CEPR London: . Discussion Paper 3341.

Bardhan, Pranab. 1984. *Land, Labor and Rural Poverty*. New York: Columbia U.P.

Barr, Abigail. 2000. "Social Capital and Technical Information Flows in the Ghanaian Manufacturing Sector." *Oxford Economic Papers* 52(3):539–59.

Barrett, Christopher B., Shane M. Sherlund & Akinwumi A. Adesina. 2004. "Macroeconomic Shocks, Human Capital and Productive Efficiency: Evidence from West African Rice Farmers." (mimeograph).

Barton, Roy F. 1969. *Ifugao Law*. Berkeley: University of California Press, 2nd edition.

Basu, Kaushik. 1986. "One Kind of Power." *Oxford Econ. Papers* 38:259–282.

Bernstein, Lisa. 1992. "Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry." *Journal of Legal Studies* XXI:115–157.

Bernstein, Lisa. 1996. "Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms." *University of Pennsylvania Law Review* 144(5):1765–1821.

Binswanger, Hans P. & John McIntire. 1987. "Behavioral and Material Determinants of Production Relations in Land-Abundant Tropical Agriculture." *Econ. Dev. Cult. Change* 36(1):73–99.

Binswanger, Hans P. & Mark R. Rosenzweig. 1986. "Behavioral and Material Determinants of Production Relations in Agriculture." *Journal of Development Studies* 22, no. 3:503–539.

Bloch, F., Garance Genicot & Debraj Ray. 2004. "Social Networks and Informal Insurance." (mimeograph).

Coate, Stephen & Martin Ravallion. 1993. "Reciprocity Without Commitment: Characterization and Performance of Informal Insurance Arrangements." *J. Dev. Econ.* 40:1–24.

Conklin, Harold. 1980. *Ethnographic Atlas of Ifugao: A Study of Environment, Culture and Society in Northern Luzon.* New Haven: Yale University Press.

Conley, T.G. 1999. "GMM Estimation with Cross-Sectional Dependence." *Journal of Econometrics* 92(1):1–45.

Conley, Timothy & Christopher Udry. 2001. "Social Learning through Networks: The Adoption of New Agricultural Technologies in Ghana." *American Journal of Agricultural Economics* 83(3):668–73.

Dekker, David, David Krackhardt & Tom Snijders. 2003. "Multicollinearity Robust QAP for Multiple-Regression." (mimeograph).

Dercon, Stefan & Joachim de Weerdt. 2002. Risk-Sharing Networks and Insurance Against Illness. Technical report CSAE Working Paper Series No. 2002-16, Department of Economics, Oxford University Oxford: .

Dercon, Stefan & Pramila Krishnan. 2000. "In Sickness and in Health: Risk-Sharing within Households in Rural Ethiopia." *Journal of Political Economy* 108(4):688–727.

de Weerdt, Joachim. 2002. Risk-Sharing and Endogenous Network Formation. Technical report WIDER Discussion Paper No. 2002-57 Helsinki: .

Ellsworth, Lynn. 1989. *Mutual Insurance and Non-Market Transactions Among Farmers in Burkina Faso.* University of Wisconsin. Unpublished Ph.D. thesis.

Evans, David. 2004. "The Spillover Impacts of Africa's Orphan Crisis." (mimeograph).

Fafchamps, Marcel. 1992. "Solidarity Networks in Pre-Industrial Societies: Rational Peasants with a Moral Economy." *Econ. Devel. Cult. Change* 41(1):147–174.

Fafchamps, Marcel. 1999. "Risk Sharing and Quasi-Credit." *Journal of International Trade and Economic Development* 8(3):257–278.

Fafchamps, Marcel. 2003. *Rural Poverty, Risk and Development.* Cheltenham (UK): Edward Elgar Publishing.

Fafchamps, Marcel. 2004. *Market Institutions in Sub-Saharan Africa.* Cambridge, Mass.: MIT Press.

Fafchamps, Marcel. 2005. "Development and Social Capital." *Journal of Development Studies* . (forthcoming).

Fafchamps, Marcel & Bart Minten. 1999. "Relationships and Traders in Madagascar." *Journal of Development Studies* 35(6):1–35.

Fafchamps, Marcel & Bart Minten. 2002. “Returns to Social Network Capital Among Traders.” *Oxford Economic Papers* 54:173–206.

Fafchamps, Marcel & Flore Gubert. 2002. “Contingent Loan Repayment in the Philippines.” (mimeograph).

Fafchamps, Marcel, Sanjeev Goyal & Marco van der Leij. 2005. “Scientific Networks and Coauthorship.” (mimeograph).

Fafchamps, Marcel & Susan Lund. 2003. “Risk Sharing Networks in Rural Philippines.” *Journal of Development Economics* 71:261–87.

Fisman, Raymond. 2003. “Ethnic Ties and the Provision of Credit: Relationship-Level Evidence from African Firms.” *Advances in Economic Analysis and Policy* 3(1) Article 4.

Foster, Andrew D. & Mark R. Rosenzweig. 1995. “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture.” *Journal of Political Economy* 103(6):1176–1209.

Foster, Andrew D. & Mark R. Rosenzweig. 2001. “Imperfect Commitment, Altruism and the Family: Evidence from Transfer Behavior in Low-Income Rural Areas.” *Review of Economics and Statistics* 83(3):389–407.

Frankenberg, Elizabeth, James P. Smith & Duncan Thomas. 2003. “Economic Shocks, Wealth, and Welfare.” *Journal of Human Resources* 38(2):230–321.

Genicot, Garance & Debraj Ray. 2003. “Group Formation in Risk-Sharing Arrangements.” *Review of Economic Studies* 70(1):87–113.

Glewwe, Paul & G. Hall. 1998. "Are Some Groups More Vulnerable to Macroeconomic Shocks than Others? Hypothesis Tests Based on Panel Data from Peru." *Journal of Development Economics* 56(1):181–206.

Good, P. 2000. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer.

Goyal, Sanjeev, Marco van der Leij & Jose Luis Moraga-Gonzalez. 2004. "Economics: An Emerging Small World?" (mimeograph).

Granovetter, M. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *Amer. J. Sociology* 91(3):481–510.

Granovetter, Mark S. 1995. *Getting a Job: A Study of Contacts and Carreers*. Chicago: University of Chicago Press. 2nd edition.

Hubert, L.J. & J. Schultz. 1976. "Quadratic Assignment as a General Data Analysis Strategy." *British Journal of Mathematical and Statistical Psychology* 29:190–241.

Johnson, Simon, John McMillan & Christopher Woodruff. 2002. "Courts and Relational Contracts." *Journal of Law, Economics, and Organization* 18(1):221–77.

Krackhardt, David. 1987. "QAP Partialling as a Test of Spuriousness." *Social Networks* 9:171–86.

Kranton, Rachel & Deborah Minehart. 2001. "A Theory of Buyer-Seller Networks." *American Economic Review* 91(3):485–508.

Krishnan, Pramila & Emanuela Sciubba. 2004. Endogenous Network Formation and Informal Institutions in Village Economies. Technical report Cambridge Working Paper in Economics No. 462 Cambridge: .

Lauby, J. & O. Stark. 1988. "Individual Migration as a Family Strategy: Young Women in the Philippines." *Population Studies* 42:473–86.

Ligon, Ethan, Jonathan P. Thomas & Tim Worrall. 2000. "Mutual Insurance, Individual Savings, and Limited Commitment." *Review of Economic Dynamics* 3(2):216–246.

Ligon, Ethan, Jonathan P. Thomas & Tim Worrall. 2001. "Informal Insurance Arrangements in Village Economies." *Review of Economic Studies* 69(1):209–44.

Lucas, Robert E.B. & Oded Stark. 1985. "Motivations to Remit: Evidence from Botswana." *J. Polit. Econ.* 93 (5):901–918.

Lund, Susan. 1996. "Informal Credit and Risk-Sharing Networks in the Rural Philippines." (unpublished Ph.D. thesis).

Meillassoux, Claude. 1971. *The Development of Indigenous Trade and Markets in West Africa.* Oxford: Oxford University Press.

Milgram, B. Lynne. 1999. Crafts, Cultivation, and Household Economies: Women's Work and Positions in Ifugao, Norther Philippines. In *Research in Economic Anthropology.* Vol. 20 Stamford, Conn.: Barry L. Isaac (ed.), JAI Press pp. 221–61.

Mitchell, J. Clyde. 1969. *Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns.* Manchester: Manchester U. P.

Mogues, Tewodaj & Michael R. Carter. 2005. "Social Capital and the Reproduction of Inequality in Socially Polarized Economies." *Journal of Economic Inequality* .

Nyblom, Jukka, Steve Borgatti, Juha Roslakka & Mikko A. Salo. 2003. "Statistical Analysis of Network Data – An Application to Diffusion of Innovation." *Social Networks* 25:175–95.

Platteau, Jean-Philippe. 1995*a*. "A Framework for the Analysis of Evolving Patron-Client Ties in Agrarian Economies." *World Development* 23(5):767–786.

Platteau, Jean-Philippe. 1995*b*. "An Indian Model of Aristocratic Patronage." *Oxford Econ. Papers* 47(4):636–662.

Quisumbing, Agnes R. 1994. "Intergenerational Transfers in Philippine Rice Villages: Gender Differences in Traditional Inheritance Customs." *Journal of Development Economics* 43(2):167–195.

Raub, Werner & Jeroen Weesie. 1990. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *Amer. J. Sociology* 96(3):626–54.

Romani, Mattia. 2003. "Love Thy Neighbour? Evidence from Ethnic Discrimination in Information Sharing within Villages in Cote d'Ivoire." *Journal of African Economies* 12(4):533–63.

Rosenzweig, M. R. & O. Stark. 1989. "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India." *Journal of Political Economy* 97(4):905–26.

Rosenzweig, Mark R. & Kenneth Wolpin. 1988. "Migration Selectivity and the Effect of Public Programs." *J. Public Econ.* 37:470–482.

Russell, Susan. 1987. "Middlemen and Moneylending: Relations of Exchange in a Highland Philippine Economy." *Journal of Anthropological Research* 43:139–61.

Shapiro, K. 1979. *Livestock Production and Marketing in the Entente States of West Africa: Summary Report*. University of Michigan.

Smith, R. & Richard Blundell. 1986. "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply." *Econometrica* 54:679–685.

Townsend, Robert M. 1994. "Risk and Insurance in Village India." *Econometrica* 62(3):539–591.

Udry, Christopher. 1994. "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria." *Rev. Econ. Stud.* 61(3):495–526.

Vega-Redondo, Fernando. 2004. "Diffusion, Search, and Play in Complex Social Networks." (mimeograph).

Weesie, Jeroen & Werner Raub. 2000. *The Management of Durable Relations.* Thela Thesis.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Mass.: MIT Press.

Table 1. Characteristics of insurance partners

A. Relationship to household head		
Close relative (*)	488	52.0%
Distant relative (**)	316	33.7%
Neighbor	104	11.1%
Friend	27	2.9%
Other (store owner, etc.)	4	0.4%
B. Residence		
Same barangay (village)	596	63.5%
Another barangay (village) in same district	151	16.1%
Elsewhere in the province	44	4.7%
Elsewhere in the Cordillera Administrative Region	35	3.7%
Elsewnere in the lowlands	58	6.2%
In Manila	12	1.3%
Abroad	14	1.5%
Unknown	29	3.1%
Total	939	100.0%

(*) son/daughter, son/daughter in law, grandchild, father/mother and brother/sister.
(**) nephew/niece, cousin, and aunt/uncle.

Table 2. Definition and mean values of variables used in regressions

		Mean	St. Dev.
Household characteristics (*n=206*)			
Location	Walking distance to the road in minutes	17.5	15.8
Sex of head	0=male, 1=female	7%	
Education of head	in years of completed education	3.9	3.4
Age of head		45.1	11.8
Health index	1=healthy, 2=frequently ill, 3=chronic illness, 4=disabled	1.403	0.8
Land inheritance	=1 if head inherited a ricefield	74%	
Value of the inheritance of the head	in 100,000 pesos	0.355	0.5
Value of the inheritance of the spouse	in 100,000 pesos	0.304	0.5
Number of children in head's family of origin		4.8	2.5
Number of children in spouse's family of origin		4.9	2.7
Birth order of household head		2.1	1.8
Head never moved	=1 if born in village	67%	
Occupation			
Occupation of head	=1 if primary occupation of head is farming	65%	
Number of members in agriculture or livestock		1.97	0.7
Number of members in casual labor		0.88	0.9
Number of members in wood carving		0.42	0.7
Number of members in carpet and basket weaving		0.22	0.5
Number of members in other self-employment activities		0.19	0.5
Number of members in salaried employment		0.23	0.5
Number of working members (*)		2.58	1.1
Total wealth	Value of fields, house, livestock and durable goods in 100,000 pesos	0.846	1.1
Total income	Earnings from jobs, crop or livestock sales and miscellenous incomes (rents, pensions, mortage payments, etc.) in 100,000 pesos	0.189	0.3
Shock variables			
Shock index in round 1	-2=Much better than 3 months ago, +2=much worse than 3 months ago	-0.752	0.6
Shock index in round 2		0.301	0.7
Shock index in round 3		-0.063	0.7
Acute sickness in round 1	=1 if at leat one member was acutely sick in last three months	26%	
Acute sickness in round 2		20%	
Acute sickness in round 3		16%	
Mild sickness in round 1	=1 if at leat one member was mildly sick in last three months	44%	
Mild sickness in round 2		38%	
Mild sickness in round 3		29%	
Unemployment in round1	=1 if at leat one member became unemployed in last three months	23%	
Unemployment in round2		20%	
Unemployment in round3		8%	
Incurred ritual expense in round 1	=1 if the household incurred ritual expense in last three months	7%	
Incurred ritual expense in round 2		10%	
Incurred ritual expense in round 3		3%	
Village dummies			
village1	=1 if household resides in village 1	26%	
village2	=1 if household resides in village 2	29%	
village3	=1 if household resides in village 3	25%	
village4	=1 if household resides in village 4	20%	
Attributes of link (*n=10,592*)			
Network link	=1 if respondent has cited j as a source of mutual insurance, 0 otherwise	2%	
Same sitio	=1 if i and j reside in same sitio	41%	
Geographical proximity	Difference in travel time (minutes) to road if i and j reside in same sitio	3.77	9.3
Correlation of i's and j's incomes		0.140	0.7

(*) The total of working member is less than the sum of members working in various activities because of double occupation.

Table 3. Regression results

Geographical proximity	coefficient estimate	dyadic t-value	coefficient estimate	dyadic t-value	coefficient estimate	dyadic t-value
Same sitio =1	2.668	**8.93**	2.662	**8.94**	2.400	**5.03**
Difference in distance to road if same sitio	-0.125	**-4.13**	-0.126	**-4.15**	-0.119	**-2.13**
Difference in:						
Dummy=1 if primary occupation of head is farming	0.000	0.00	-0.004	-0.03	0.087	0.48
Number of working members	0.025	0.43	0.023	0.32	0.074	0.79
Age of household head	-0.010	**-2.72**	-0.010	**-2.78**	-0.009	-1.64
Health index 1-4 (1=good health, 4=disabled)	0.035	0.60	0.035	0.69	0.050	0.79
Years of education of household head	-0.012	-0.66	-0.012	-0.72	-0.007	-0.24
Total wealth (predicted)	-0.125	**-2.62**	-0.125	**-2.56**	-0.209	**-2.49**
Sum of:						
Dummy=1 if primary occupation of head is farming			0.013	0.11	0.107	0.50
Number of working members			-0.157	**-2.56**	-0.137	-1.55
Age of household head			0.002	0.33	0.010	0.99
Health index 1-4 (1=good health, 4=disabled)			0.102	1.24	0.179	1.20
Years of education of household head			0.005	0.26	0.037	0.96
Total wealth (predicted)			0.014	0.15	0.030	0.18
Village dummies:			included but not shown			
Mills ratio					-0.116	-0.10
intercept	-5.828	-16.93	-5.538	-8.37	-6.664	-5.01
Number of observations	10592		10592		4788	

Note: the dependent variable=1 if i cites j as source of mutual insurance, 0 otherwise
Estimator is logit. All t-values based on standard errors corrected for dyadic correlation of errors.

Table 4. With a more detailed breakdown of income earning activities

Geographical proximity	coefficient estimate	dyadic t-value	coefficient estimate	dyadic t-value
Same sitio =1	2.669	**8.93**	2.676	**8.99**
Difference in distance to road if same sitio	-0.125	**-4.13**	-0.128	**-4.20**
Difference in:				
Nber of members in agriculture or livestock	-0.118	-1.32	-0.127	-1.38
Nber of members in casual labor	-0.002	-0.05	-0.009	-0.16
Nber of members in wood carving	0.064	0.68	0.055	0.60
Nber of members in carpet and basket weaving	-0.135	-0.97	-0.123	-0.92
Nber of members in other self-employed activity	-0.117	-1.05	-0.145	-1.12
Nber of members in salaried employment	-0.071	-0.49	-0.076	-0.50
Years of education of household head	-0.008	-0.50	-0.010	-0.65
Age of household head	-0.009	**-2.44**	-0.009	**-2.48**
Health index 1-4 (1=good health, 4=disabled)	0.027	0.48	0.026	0.55
Number of working members	0.062	0.76	0.062	0.71
Total wealth (predicted)	-0.110	**-2.15**	-0.109	**-2.23**
Sum of:				
Nber of members in agriculture or livestock			-0.123	-1.53
Nber of members in casual labor			-0.086	-0.96
Nber of members in wood carving			0.036	0.29
Nber of members in carpet and basket weaving			0.067	0.46
Nber of members in other self-employed activity			-0.114	-0.90
Nber of members in salaried employment			-0.004	-0.03
Years of education of household head			-0.002	-0.10
Age of household head			0.003	0.43
Health index 1-4 (1=good health, 4=disabled)			0.093	1.12
Number of working members			-0.111	-1.29
Total wealth (predicted)			0.024	0.24
Village dummies:		included but not shown		
intercept	-5.857	-16.45	-5.214	-6.50
Number of observations	10592		10592	

Note: the dependent variable=1 if i cites j as source of mutual insurance, 0 otherwise
Estimator is logit. All t-values based on standard errors corrected for dyadic correlation of errors.

Table 5. With income correlation, mean, and standard deviation

Geographical proximity	coefficient estimate	dyadic t-value	coefficient estimate	dyadic t-value
Same sitio =1	2.647	**8.84**	2.655	**8.84**
Difference in distance to road if same sitio	-0.121	**-3.90**	-0.121	**-3.92**
Incomes variables (instrumented)				
Correlation of i and j's incomes	1.083	1.44	1.370	1.41
Sum of i and j's mean incomes			0.216	0.21
Difference in i and j's mean incomes			0.734	0.76
Sum of i and j's income standard deviation			-7.365	**-2.52**
Difference in i and j's income standard deviation			-9.086	**-2.31**
Difference in:				
Dummy=1 if primary occupation of head is farming	0.028	0.23	0.088	0.76
Number of working members	0.003	0.06	-0.059	-0.95
Age of household head	-0.010	**-2.52**	-0.009	**-2.24**
Health index 1-4 (1=good health, 4=disabled)	0.027	0.46	-0.057	-0.91
Years of education of household head	-0.010	-0.59	0.007	0.31
Total wealth (predicted)	-0.113	**-2.37**	0.024	0.29
Village dummies:		included but not shown		
Intercept	-5.995	-15.41	-5.464	-10.71
Number of observations	10264		10264	

Note: the dependent variable=1 if i cites j as source of mutual insurance, 0 otherwise
Estimator is logit. All t-values based on standard errors corrected for dyadic correlation of errors.

Table 6. Gifts, loans and networks

	Gifts received (*)		Loans received (*)	
Geographical proximity	**coefficient estimate**	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**
Same sitio =1	0.052	**6.51**	0.037	**4.17**
Difference in distance to road if same sitio	-0.002	**-5.27**	-0.001	**-2.55**
Network dummy x difference in:				
Dummy=1 if primary occupation of head is farming	-0.234	-1.25	-0.255	**-2.32**
Number of working members	-0.080	-0.72	-0.003	-0.07
Age of household head	0.005	0.61	0.001	0.18
Health index 1-4 (1=good health, 4=disabled)	0.281	**2.62**	0.011	0.13
Years of education of household head	-0.018	-0.66	-0.009	-0.47
Total wealth (predicted)	-0.034	-0.28	0.019	0.26
Difference in:				
Dummy=1 if primary occupation of head is farming	0.001	0.51	-0.009	**-2.36**
Number of working members	0.000	1.38	0.002	1.08
Age of household head	0.000	-0.44	0.000	-0.52
Health index 1-4 (1=good health, 4=disabled)	0.001	**2.04**	0.000	-0.22
Years of education of household head	0.000	-0.80	0.000	-0.74
Total wealth (predicted)	0.000	0.49	-0.002	-0.71
Sum of:				
Dummy=1 if primary occupation of head is farming	0.003	0.99	-0.002	-0.50
Number of working members x number of activities	-0.003	-1.52	0.000	0.13
Age of household head	0.000	1.13	0.000	-0.76
Health index 1-4 (1=good health, 4=disabled)	0.003	1.15	0.001	0.64
Years of education of household head	0.000	0.35	0.001	**2.20**
Total wealth (predicted)	0.001	0.63	-0.001	-0.39
Village x time dummies	included but not shown			
Intercept	-0.015	-0.83	-0.006	-0.28
Number of observations	21184		21184	

Estimator is least squares. All t-values based on standard errors corrected for dyadic correlation of errors.
(*) Dependent variable in log(value of gift or loan +1)

Table 7. The insurance role of gifts and loans between villagers

	Gifts received (*)		Loans received (*)	
Insurance	**coefficient estimate**	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**
Difference in current status (predicted)	0.000	-0.03	0.009	1.10
Network dummy x difference in current status	0.613	**1.71**	0.126	0.41
Network dummy	0.996	**10.54**	0.370	**4.88**
Geographical proximity				
Same sitio =1	0.010	**2.54**	0.022	**2.95**
Difference in distance to road if same sitio	-0.001	**-3.20**	-0.001	**-1.63**
Difference in:				
Dummy=1 if primary occupation of head is farming	-0.003	-0.60	-0.014	**-3.23**
Number of working members	-0.002	-1.06	0.001	0.79
Age of household head	0.000	1.51	0.000	0.23
Health index 1-4 (1=good health, 4=disabled)	0.006	**2.41**	-0.001	-0.30
Years of education of household head	0.000	-0.50	0.000	-0.65
Total wealth (predicted)	0.002	0.91	-0.001	-0.27
Village x time dummies	included but not shown			
intercept	-0.002	-0.43	-0.006	-1.47
Number of observations	21184		21184	

Estimator is least squares. All t-values based on standard errors corrected for dyadic correlation of errors.
(*) Dependent variable in log(value of gift or loan +1)

Table A1. Instrumenting wealth

	Sum of wealth		Wealth difference	
	coefficient estimate	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**
Regressors	**all regressors as sums**		**all regressors as diff.**	
Dummy=1 if born in village	0.080	0.92	0.080	0.89
Dummy=1 if head is male, 2 if female	-0.305	**-1.94**	-0.305	**-1.87**
Education of household head	0.034	**1.89**	0.034	**1.82**
Number of siblings of household head	-0.009	-0.46	-0.008	-0.44
Number of siblings of spouse	-0.008	-0.62	-0.008	-0.60
Value of inherited land of head	0.940	**6.99**	0.941	**6.74**
Value of inherited land of spouse	1.028	**5.09**	1.028	**4.92**
Village dummies		included but not shown		
Intercept	1.019	1.94	n.a.	
Number of observations	10592		10592	

Estimator is least squares. All t-values based on standard errors corrected for dyadic correlation of errors.
(*) Dependent variable in log(value of gift or loan +1)

Table A2. Migration selection regression

Sum of:	coefficient estimate	dyadic t-value
Birth order of household head	-0.118	**-1.83**
Dummy=1 if male head, 2 if female	-0.165	-0.75
Education of household head	-0.055	**-1.91**
Age of household head	-0.003	-0.43
Village dummies	included	
intercept	1.561	1.59
Number of observations	10592	

Dependent variable=1 if both household heads reside in birth village
Estimator is logit.
All t-values based on standard errors corrected for dyadic correlation of errors.

Table A3. Instrumenting regressions for income correlation, mean and standard deviation

	Correlation of income		Sum of mean income		Sum of income std. dev.		Difference in mean income		Difference in income std. dev.	
Geographical proximity	**coefficient estimate**	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**	**coefficient estimate**	**dyadic t-value**
Same sitio =1	-0.014	-0.59	-0.001	-0.07	-0.007	-0.99				
Difference in distance to road if same sitio	-0.002	-0.97	-0.001	-1.64	0.000	-0.41				
Sum of:										
Nber of members in agriculture or livestock	0.021	0.97	-0.048	-0.83	0.011	1.31				
Nber of members in casual labor	0.005	0.26	-0.025	-0.81	0.004	0.89				
Nber of members in wood carving	0.041	1.17	-0.013	-0.32	0.000	0.07				
Nber of members in carpet and basket weaving	0.045	1.06	0.023	0.45	0.027	1.47				
Nber of members in other self-employed activity	0.003	0.14	-0.039	-1.34	0.005	0.41				
Nber of members in salaried employment	0.071	1.43	-0.091	**-2.25**	0.003	0.19				
Number of working members	-0.072	**-3.62**	0.073	**1.83**	-0.010	**-1.72**				
Total wealth	0.010	0.43	0.067	**3.15**	0.024	1.43				
Years of education of household head	0.003	0.59	-0.002	-0.35	0.001	0.29				
Health index 1-4 (1=good health, 4=disabled)	0.035	1.64	-0.024	-1.24	-0.011	**-1.93**				
Difference in:	**absolute difference**						**difference**		**difference**	
Nber of members in agriculture or livestock	0.066	**2.97**					-0.045	-0.74	0.012	1.29
Nber of members in casual labor	-0.026	-1.29					-0.023	-0.75	0.004	0.81
Nber of members in wood carving	0.005	0.15					-0.012	-0.30	0.000	-0.01
Nber of members in carpet and basket weaving	0.016	0.49					0.096	1.19	0.043	**1.78**
Nber of members in other self-employed activity	0.018	**3.68**					-0.088	-1.61	-0.006	-0.31
Nber of members in salaried employment	0.010	0.59					-0.153	**-2.09**	-0.014	-0.68
Number of working members	-0.014	-0.87					0.058	**1.74**	-0.011	-1.58
Total wealth	-0.029	-1.14					0.079	**3.26**	0.027	1.47
Years of education of household head	0.009	**2.31**					0.020	0.89	0.006	1.09
Health index 1-4 (1=good health, 4=disabled)	0.002	0.26					-0.022	-1.26	-0.012	**-1.93**
Village dummies						included but not shown				
intercept	-0.189	-1.32	0.465	1.23	0.081	2.20				
Number of observations	10264		10486		10486		10592		10592	

Estimator is least squares. All t-values based on standard errors corrected for dyadic correlation of errors.

Table A4. Instrumenting regression for shock index

Difference in:	coefficient estimate	dyadic t-value
Dummy=1 if acute sickness	0.415	**3.62**
Dummy=1 if mild sickness	0.205	**2.84**
Dummy=1 if one household member unemployed	0.219	**2.22**
Dummy=1 if incurred ritual expense	0.562	**4.13**
Number of observations	21184	

Dependent variable is difference in shock index between i and j
Estimator is least squares.
All t-values based on standard errors corrected for dyadic correlation of errors.