VILLAGE ECONOMIES AND THE STRUCTURE OF EXTENDED FAMILY NETWORKS[*]

MANUELA ANGELUCCI[†]   GIACOMO DE GIORGI[‡]

MARCOS A.RANGEL[§]   IMRAN RASUL[¶]

APRIL 2007

Abstract

We provide evidence from rural Mexico on how the characteristics of village economies relate to the structure of extended family networks that reside within them. Our analysis sheds light on how the overall level of economic development of the village, and the degree of village inequality, relate to the number and size of family networks in the village, as well as the precise inter and intra-generational family links that bind members of each network. We find that in more marginal villages, and villages in which households are equally poor, there are significantly fewer family dynasties present. Each dynasty is of larger size both in terms of the absolute number of households, and the share of village households that are embedded within the network. The reason why family networks are larger in such villages is that heads of household and their spouses are both significantly more likely to have their siblings and adult sons and daughters resident within the village. Finally, we find that controlling for a rich set of household characteristics, these features of the village economy continue to shape family networks over and above the characteristics of individual households within them.

Keywords: extended family network, village inequality, village marginality.

JEL Classification: J12, O12, O17.

1

1 Introduction

Village economies in developing countries are typically characterized by missing or imperfect markets, the absence of formal institutions of contract enforcement, and low levels of infrastructure and other forms of public goods provision. These features of the local economy serve both to constrain the economic choices of the poor, and leads them to endogenously develop strategies to make such constraints less binding. In this paper we focus on one such mechanism that is in principle available to all households, namely choices related to the physical presence and characteristics of their extended family.[1]

We provide evidence from rural Mexico on how the characteristics of village economies relate to the structure of extended family networks that reside within them. Our analysis sheds light on how the overall level of economic development of the village, and the degree of village inequality in income, assets, and land, relate to the number and size of family networks in the village, as well as the precise inter and intra-generational family links that bind members of each network.

While undoubtedly other households outside of the extended family network also influence household behavior in village economies, there are good reasons to focus on family networks. First, evolutionary biology suggests preferences are defined over the family dynasty. Moreover there are specific inter-generational investments – such as those into children's education, bequests, and the choice of marriage partners – that have no counterpart in relationships between friends, and provide long run incentives for family members to reciprocate in implicit contractual arrangements. Finally, the transactions costs of such arrangements with non-family members may be higher because it is both more costly to observe outcomes outside the family network, and fewer mechanisms exist by which to punish non-family members that renege on such agreements [La Ferrara 2003].

To document the relationship between the village economy and family network structure we use household data from the evaluation of *Progresa*, a social assistance program in rural Mexico. This data has two key features that we exploit. First, we combine information on the paternal and maternal surnames of heads of households and their spouses, with the patronymic naming convention in Mexico to build two types of extended family link outside the household but within the village – (i) inter-generational links, such as those from the head and spouse to their parents, and to their adult sons and daughters; (ii) intra-generational links, such as those from the head and spouse to their brothers and sisters. We combine this with information from the household roster to identify extended family members that co-reside in the household. These sources of

---

[1]Duflo and Banerjee [2007] review data from 13 countries to document the economic lives of the poor in light of them facing imperfect markets for credit, insurance, savings, labor, and land, and the strategies used to deal with such constraints. The notion that there may exist an interplay between local family networks and the economic environment poor households face has been well recognized in the literature [Bardhan 1984, Basu 1986].

information together provide an almost complete mapping of extended family structures across 500 villages in rural Mexico, covering around 22,000 households and over 130,000 individuals.

Second, we exploit household and village level information to construct data on two key aspects of the village economy. The first is an index of the marginality of the village. This is specifically designed by the Mexican federal government to measure the extent of local development. The second is an index of each household's permanent income, from which we construct the standard deviation of permanent income in the village. This captures the extent of within village inequality in the distribution of income, assets, and land. These two indices capture long run features of the economic environment in each village.[2]

Core to our analysis is relating these features to the extended family networks observed within the village. A number of mechanisms exist for why the overall level of development of the village economy and family network structures should be closely related.

The first relates to risk pooling arrangements across households. While such *ex post* risk pooling mechanisms have been documented to exist in a number of developing country settings, tests of the Pareto efficient allocation of risk being achieved at the village level are typically rejected [Townsend 1994, Ligon 1998, Dercon and Krishnan 2000]. However there is stronger evidence of risk pooling within ethnic groups [Deaton 1992, Udry 1994], sub-castes [Munshi and Rosenzweig 2005], and family and friends [Rosenzweig 1988, Fafchamps and Lund 2003, La Ferrara 2003, Cox and Fafchamps 2007]. There is also evidence that interpersonal loans, transfers, and labor exchange arrangements occur primarily within long-standing interpersonal networks such as the extended family [Ellsworth 1985, Lucas and Stark 1985, Krishnan and Sciubba 2004].

Other mechanisms relating levels of development of the village economy to family networks include the choice of marriage partners for children, and in particular daughters, as a means of insurance [Rosenzweig and Stark 1989]; the inheritance of land, or the inter-generational transmission of land specific human capital, and the presence of adult children, in particular sons [Foster 1993]; and the nature of household production and joint consumption of family public goods [Foster and Rosenzweig 2002]. Moreover, changes in local economic opportunities can also trigger household partition and migration [Foster 1993].

There also exist mechanisms relating village level inequality and family network structures. For example, the gains from informal risk sharing contracts increase as the incomes of those party to such arrangement become less correlated. Inequality within a village may therefore correspond to households having different income profiles and being subject to different sources of risk. On the one hand, this implies family networks should not be geographically concentrated, as they will

---

[2]The marginality index is measured as of 1995, and the index based on measures of permanent income is defined as of 1997.

3

then be more likely to be subject to common climatic related risks. On the other hand, as the geographic dispersion between network members increases, the transactions costs of maintaining interpersonal links, monitoring the behavior of others in the network, and providing non-monetary transfers such as labor exchange, also increase.

Marriage markets provide an important mechanism through which village inequality and family network structures relate to each other. If matching in the marriage market is assortative in terms of family resources, inequality leads to less integrated villages as households marry within wealth classes. Inequality within villages then becomes self perpetuating as family networks do integrate with each other over time [Kremer 1997, Fernández *et al* 2005, Ermisch *et al* 2006].

To the extent that village inequality generates distinct group identities, this can affect the accumulation of social capital and provision of local public goods [La Ferrara 2003]. Theoretically, the effect of inequality on public goods provision is ambiguous and depends on the characteristics of the public good. With increasing returns to scale due to large fixed costs, inequality can favour collective action [Olson 1965]. However, the presence of market imperfections and decreasing returns to productive assets, can generate the opposite result [Bardhan *et al* 2002]. Moreover, to the extent that the distribution of assets determines the level of control each member has over the common good, inequality might favour rent seeking and reduce efficiency [Banerjee *et al* 2001].

To summarize, there are a multitude of mechanisms through which the marginality of the village economy and inequality within it, might relate to the family network structures that are observed within the village. The theoretical predictions on how these factors interplay with each other are not clear cut, and therefore the relationships between them remains an open empirical question. This paper provides a first attempt to document these correlations.

We present five main findings. First, in terms of the overall structure of village networks, we find that in more marginal villages there are significantly fewer family dynasties. Each dynasty is of larger size both in terms of the absolute number of households, and the share of village households that are embedded within the network. Second, in more unequal villages, there are significantly more family dynasties, and each encompasses fewer households within it.

In short, in more marginal villages, and in which households are equally poor, we observe the existence of fewer distinct family dynasties, and each household being embedded within a larger extended family network.

Third, in terms of the specific family links that exist in the network, we find the reason why family networks are larger in more marginal and more equal villages, is that heads of household and their spouses are both significantly more likely to have their siblings to be resident within the same village. This applies equally to brothers and sisters of heads and their spouses.

Fourth, in terms of the relation between the presence of inter-generational links and the village

4

economy, the presence of adult daughters within the household is particularly sensitive to characteristics of the village economy. In particular, in more marginal and more equal villages, adult daughters are significantly less likely to reside *within* the household, and significantly more likely to reside in *another* household within the same village.

Fifth, moving from a village level analysis to a household level analysis, we find these results to be replicated once a rich set of household characteristics are controlled for. In particular, in more marginal or more equal villages, adult sons are significantly more likely to reside both within the village or be coresident within the household. In contrast, while in more marginal or more equal villages daughters are significantly more likely to reside within the village, these features lead them to be *less* likely to remain coresident within the household. These results highlight that the structure of family networks relates to features of the village economy as a whole, over and above the characteristics of individual households within them.

As these relations are identified from a cross section of villages and households, then absent any exogenous variation in the marginality of the village or inequality within it, it is impossible to identify whether network structures endogenously respond to these characteristics of the village economy, or whether the local economy is itself shaped by the nature of extended family networks that reside within it. In either case, the results show that any attempt to analyze the cause and effect of why villages differ in their levels of poverty and inequality, cannot be done in isolation from an analysis of why and how extended family networks also differ across villages.

The remainder of the paper is organized as follows. Section 2 describes the data, the economic characteristics of the villages we study, and the algorithm used to construct extended family ties. Section 3 provides descriptive evidence on the extended family structures at the village level, network level, and household level. Section 4 provides empirical evidence on the relationship between characteristics of the village economy on the overall number of family dynasties, the structure of these networks, and the specific types of inter and intra-generational family link that bind together households within them. Section 5 concludes.

2 The Data and Construction of Extended Family Links

2.1 The Progresa Evaluation Data

We use household data collected to evaluate the *Progresa* social assistance program in rural Mexico. The program provides grants to households in the form of conditional cash transfers related to their children's school attendance, and attendance at local health clinics. Household data was collected on a panel of around 24,000 households every six months in 500 villages between March 1998

and November 1999. Within each village, a full census of households is surveyed. The villages are located in seven out of thirty one states in Mexico. These states are Guerrero, Hidalgo, Michoacan, Puebla, Queretaro, San Luis Potosi, and Veracruz, and are mainly located in the south of Mexico.

Villages were selected on the basis of a village marginality index constructed from information on – (i) the share of illiterate adults in the village; (ii) the share of dwellings without water, drainage systems, electricity, and with floors of dirt; (iii) the average number of occupants per room in households; (iv) the share of the population working in the primary sector; (v) health and school infrastructures present in the village; (vi) distances from other villages. The village marginality index therefore provides a summary statistic characterizing average levels of poverty within the village, the vulnerability of the local population to aggregate income shocks, and the integration of the village economy into local markets and the wider economy. The index is defined so that it increases as the village becomes more marginal or worse off.

We restrict our analysis to the 22,553 households that are surveyed in the first and third waves of the data. As we aim to compare and contrast the extended family links of heads and their spouses, we focus attention on the 85% of households that are couple headed, where the head is always defined to be male.[3]

To determine a household's entitlement to *Progresa* transfers, in 1997 households were classified as either being eligible (poor) or non-eligible (non-poor) according to a household welfare index. This index is calculated as a weighted average of – (i) household income (excluding children); (ii) household size; (ii) ownership of durables, land, and livestock; (iv) educational attainment; (v) physical characteristics of the dwelling. The index is designed to give relatively greater weight to the correlates of permanent, rather than current, income.[4]

We use this household welfare index to construct a measure of inequality at the village level – defined to be the standard deviation of the welfare index across all households in the village. This summary statistics captures inequality in income, acquired human capital, assets, and land holdings, across a census of households in the village.[5]

In order to more easily interpret the effect of village marginality on family networks, we stan-

---

[3]Of the 506 villages, 320 were randomly assigned to the treatment group, namely locations where *Progresa* would later be implemented, and 186 villages were assigned to be control villages. In particular, *Progresa* transfers began to be distributed in May 1998, between the second and third waves of data. We do not exploit any of the randomized design features of the program in this paper. The bulk of our analysis uses data collected in waves three or before. This allows us to present evidence on the pre-program relationship between features of the village economy, and the structure of extended family networks, without confounding the relationship with any direct effects of the program on extended family networks.

[4]In October 1997, 52% of households were classified to be eligible for *Progresa* according to this welfare index.

[5]We later complement this data with information from community surveys designed to measure features of the village economy that are not explicitly captured by either the marginality index of the village inequality index. This includes information on the prevalence of natural disasters that have afflicted the village which may relate to aggregate levels of uncertainty or the persistence of shocks faced by households in the village.

6

dardize the marginality index across all villages so that it has mean zero and a standard deviation of one. The household welfare index, on which the village inequality index is based, is defined separately for each of the seven states covered in the *Progresa* data. We therefore standardize the village inequality index by state, so that it has mean zero and a standard deviation of one in each state. Table 1 presents evidence on these village level measures of the economic environment, averaged across all villages in the sample. The table also presents information on the average village size, as defined by the number of households in the village.

The sample villages are relatively small so that family networks may be important determinants of household behavior and outcomes in this empirical setting. A useful stepping stone is therefore to establish the correlates of family networks in the first place.

Figure 1A shows a scatter plot of the village marginality and inequality indices. This shows the extent of across village level variation in both indices, and highlights that the two are negatively correlated – unconditionally, more marginal villages are associated with lower levels of economic inequality. The figure also shows there exists a wide range of inequality across villages at each and every level of village marginality – namely, for any given level of village marginality, there are some villages in which households are equally poor, and in others in which households have very dispersed levels of permanent income. This opens up the possibility of separately identifying the effects of each channel on the structure and composition of extended family networks.

Finally, Figure 1B shows that the two measures of the village economy are not much related to village size. The correlation in the village marginal (inequality) index and village size is -.151 (.028). This is important as there are some features of network structures that are mechanically correlated to village size. The fact that this is orthogonal to our measures of the village economy eases concerns that we spuriously attribute any mechanical correlation between village size and networks to these characteristics of the economic environment.

2.2 Constructing Extended Family Links

2.2.1 Surnames and the Matching Algorithm

To identify the extended family links between households in the same village we exploit information on surnames provided in the third wave of data. We combine this information with the patronymic naming convention in Mexico to build two types of family link – (i) intra-generational family links, such as those from the head (spouse) of the household to his (her) brothers and sisters; (ii) inter-generational family links, such as those from the head (spouse) of a household to his (her) parents, and from the head and spouse of the household to their adult sons and daughters.[6]

---

[6]Two concerns arise from the surnames data being measured in the first wave of post-program data. First, households may endogenously respond to the program by changing household structures, in particular, by artificially

Mexicans use *two* surnames – the first is inherited from the father's paternal lineage and the second from the mother's paternal lineage. For example, former Mexican president Vicente Fox Quesada would be identified by his given name (Vicente), his father's paternal name (Fox) and his mother's paternal name (Quesada). In the evaluation data, respondents were asked to provide the – (i) given name; (ii) paternal surname; and, (iii) maternal surname, for each household member. Hence couple headed households have four associated surnames – the paternal and maternal surnames of the head, and the paternal and maternal surnames of his wife.[7]

Figure 2 provides an illustration of the matching algorithm. To define each family link, we use information on *two* of the four surnames. Consider household **A** at the root of the family tree. The head of the household has paternal and maternal surnames $F1$ and $f1$ respectively. His wife has paternal and maternal surnames $F2$ and $f2$ respectively.[8]

The children of the couple in household **A** will adopt the paternal surnames of their father ($F1$) and mother ($F2$). Hence we define there to be a parent-son relationship between households **A** and **B** if – (i) the paternal surname of the head in household **B** is the same as the paternal surname of the head in household **A** ($F1$); and, (ii) the maternal surname of the head in household **B** is the same as the paternal surname of the spouse in household **A** ($F2$). Parent-daughter relationships can be similarly defined. Moreover, intra-generational family ties between siblings can also be identified. For example, the heads of households **B** and **C** are identified to be brothers if they share the same paternal and maternal surnames.

In Figure 2 we assume households are couple headed solely to ease the exposition. To deal with the 15% of households that are single headed we use information on the gender of the head

---

forming new households in order to increase the number of eligibles in the family. This concern is ameliorated by the fact that the register of eligible households was drawn up at baseline, and only households recorded to be eligible at that point were later entitled to receive transfers. Moreover, although there is an increase in the number of households from the baseline to October 1998, this increase is proportionately the same in both treatment and control villages. A second concern is that the program may affect the migration of the household head or of his spouse. However, only .4% (.5%) of households in wave 3 (5) report having a migrant head or spouse.

[7]The precise wording of the question in Spanish is, "*Dígame por favor el nombre completo con todo y apellidos de todas las personas que viven en este hogar, empezando por (jefe del hogar) – (i) nombre; (ii) apellido paterno; (iii) apellido materno*". We cleaned the surnames data as follows – (i) we removed non-alphabetical characters, replaced "Sin Apellido" (no surname) with missing values, and corrected some obvious typos based on intra-household surname checks; (ii) we imputed a small number of missing female surnames from wave 2; (iii) we verified surnames using the same information from wave 5, and verified the relationship to the household head using wave 1 data. No information on surnames is available in the first wave of data. The head of household is originally defined to be the main income earner. In a very small number of cases the head of a couple headed household is reported to be a women. To keep clear the exposition, we redefine the head to be male in such cases.

[8]We use the convention that the head's surnames are written in black, and those of his wife are written in red italics. Paternal surnames are indicated in upper case and maternal surnames are indicated in lower case. First names are not shown as they are not relevant for the construction of extended family links. Each household in the family tree is assumed to be couple headed purely to ease the exposition. In Anglo Saxon countries, $F1$ corresponds to the family name and $F2$ corresponds to the spouse's maiden name.

to accurately define each family link. Finally, we impose the following restrictions when defining family links – (i) inter-generational links exist when the relevant individuals have at least 15 years age difference, and no more than 60 years age difference between mother and child; (ii) intra-generational links exist when the individuals have at most 30 years age difference.

However, there are limits to which information on surnames can be used to construct family ties. Consider links from household i to a single headed household j. As Figure 1 shows, the fact that household j is single headed does not affect the construction of links from the head and spouse of household i either to their children or to their siblings. However, links from the head (spouse) of household i to the household of his (her) parents can only be identified if *both* his (her) parents are alive and resident together. This is because this particular family link is identified using information from household j on the paternal surnames of both the head and spouse.[9]

2.2.2 Measurement Error in Extended Family Links

There are a number of potential forms of measurement error in the surnames data that can be checked for. The first arises from the convention that women change their paternal surname to their husband's paternal surname at the time of marriage. To address this concern, we note that the precise wording of the question specifically asks respondents to name the paternal and maternal surname of each household member. Furthermore, in only 5.8% of households is the spouse's maternal surname recorded to be the same as her husband's paternal surname. This provides an upper bound on the extent to which measurement error of this form is occurring.

Second, if the male head is the respondent, he may not recall his wife's maternal surname and simply replace it with her paternal surname. This may occur because his children only inherit his wife's paternal surname. Reassuringly, this problem occurs in only 4.9% of households. A final circumspect case is households in which the paternal and maternal surnames of both the head and spouse are all reported to be the same. This occurs for 1.6% of households, although the figure drops to .5% if we exclude households with the most common surname in the data.

Some forms of measurement error however cannot be addressed. The first arises from any remaining typos in surnames. Second, there may be two identical families in the village who share the same paternal and maternal surnames of head and spouse but are genuinely unrelated. The matching algorithm then assigns the number of family links to be double what they actually are. A check for the severity of this problem is based on the following intuition. By definition, household i cannot have parental links to more than two other households (the parent's of the head and

---

[9]However this is unlikely to be a major issue. For example we note that female widows aged above 40 are 37% more likely to live as a dependent within a household, rather than head their own household, relative to a similar married woman. These single parents are then recorded in the household roster.

the parent's of the spouse), conditional on the parents not being present within the household. This is true for 97% of households using our matching algorithm. Third, consider a scenario in which a women's brother marries someone with the same maternal surname as himself. Then the woman's niece will be identified as her sister and although the households are within the same family network, the strength of their tie may be inferred to be stronger than it actually is.

With these caveats in mind, it is equally important to reiterate that all links are defined across households on the basis of *two* surname matches and the fact that these two surnames map precisely to two of the four surname types. For the empirical analysis, we only exploit information on whether such family links exist for a household in the village, and not the number of links.

3   Descriptive Evidence

3.1   Village Characteristics

Table 2 reports the share of households within a village level that have each type of family link. The first point to note is that 22% of couple headed households are isolated in that they have no extended family members resident in the village, and the remaining 78% of households are connected in that they have at least one of their extended family members living within the same village. The incidence of not having extended family members geographically proximate in the same village is therefore at least as high as the incidence of single headedness, which affects 15% of households.

The remaining columns split family links into – (i) inter-generational links to parents and adult children who head their own households; (ii) intra-generational links to siblings who head their own households. We see that 16.4% of households have ties between the head of household and that of his parents – this is around double the number of parental links from the spouse to her parents in the village. The same inter-generational pattern occurs when we consider links from the head and spouse to their adult children – households are around twice as likely to have at least one son resident in the village than at least one daughter.

Intra-generational ties are more prevalent than inter-generational ties. The most likely type of extended family tie to exist is that between two heads of household, namely brothers – 45.2% of all households are such that the head of household has at least one brother in the village. In contrast, only 23% of households have ties from the spouse to her sister.

3.2 The Number of Extended Family Links

Table 3 provides details on the number of specific family links each household has to others in the village. These are reported for connected households, namely those households that are embedded within an extended family network in the village and so have at least one family member resident in the village. As we focus on couple headed households, we are able to report each type of link from the head and spouse separately.

The first column shows that the number of parents present outside the household and within the village is higher for the head than for his spouse. The number of links to adult children are, by construction, identical for head and spouse. On average, respondents have .443 sons that head their own household in the village, and .209 daughters. In terms of sibling links, heads again have more links to their siblings than do their spouses. In total, heads of household have blood ties to 3.34 others in the same village, and their spouse is connected to 2.54 other households. Given that by construction links to adult sons and daughters are the same for heads and their spouses, this implies that on average, a household is directly connected to 5.23 other households in the village.

Four further points are of note. First, Table 3 decomposes the variation in each statistic into that between and within villages. We see that while in most cases there is more variation within than between villages, the proportion of the variation that is attributable to cross village differences tends to be higher for female defined links than for male links.

Second, we can also use the household roster to measure family ties that are coresident. When doing so we note that parents are significantly more likely to reside outside of the household than within it. On average, the head (spouse) has .06 (.02) parents resident with him (her). In other words, there are at least ten times as many parents outside the household and within the village, than are coresident with respondents. Given respondents' ages, there are many more young children coresident inside the household than adult children outside the household. Sibling links are also many times more likely to be outside the household, which is as expected given the age of respondents. Heads have more siblings links than their spouses, and this is again true for links both inside and outside the household.[10]

Third, the evidence in Table 3 shows that the majority of extended family links are those of male heads of household. A key source of this difference arises from women moving households and villages at the time of marriage. To shed light on this, we exploit data on spouse's marital history. Wives were asked about where they went to live after marriage – 49.3% stated that they went to live with their in-laws after marriage, and only 6.5% report living with their own parents. The key difference between spouses with and without parents resident in the village is that 85% of spouses that have their parents present in the village report remaining in the same village at

---

[10]In October 1998 the average age of heads (spouses) among couple headed households is 45.0 (40.5).

the time of marriage. The figure for spouses that have no parental links in the village is only 61%. In line with the evidence from rural Indian in Rosenzweig and Stark [1989], these descriptives hint that in rural Mexico, parents may choose to marry their daughters into distant villages as a consumption smoothing device.

Finally, to provide external validity to this constructed data, in Angelucci *et al* [2007] we present similar information from an alternative data source that was collected in a comparable economic environment and time period. The *Mexican family Life Survey* (MxFLS), collected in 2001, provides information on the numbers of each type of family link, by head and spouse, that are alive in *any* location, not just the same village. The MxFLS therefore provides upper bounds on what should be recorded as family links in the *Progresa* data. Reassuringly, we find that in the MxFLS, the number of family links to parents, adult children, and siblings outside the household and located anywhere, are indeed greater than those we construct within the village.

3.3   Extended Family Networks

Having described the number of family links that connected households have to others in the same village, we now describe the extended family network as a whole. We use tools from the analysis of social networks to describe the family networks in the *Progresa* villages [Wasserman and Faust 1994]. To understand when two households are defined to be in the same family network, consider a scenario in which the heads of households i and j are linked because they are brothers. Household j may itself be linked to household k, say, because the parents of the spouse of household j reside in household k. Households i and k then lie within the same family network, even though they do not have a *direct* family link between them. Following the terminology in , households i and j are said to be of distance one from each other and households i and k are of distance two from each other. Two households i and j are then defined to be within the same family network if the distance between them, d_{ij}, is finite.

For this analysis we do not consider households that are single nodes – namely those unconnected to any other household in their village. We therefore consider family networks with at least two members. There are 2196 such family networks, covering 17,030 households.[11]

To begin with, Figure 3A shows the number of family networks in a village. There are around five to ten different family dynasties within the same village. Figure 3B shows the size distribution

---

[11]Single nodes correspond to three types of household. First, those in which all family relations are living under the same roof. Second, there may exist households for whom there are no children or siblings of the head or spouse heading their own households in the village, and for whom the parents of the head and/or spouse are outside of the household but are widowed. Finally, there may be genuine cases of households for whom there are no children or siblings of the head or spouse heading their own households in the village, and for whom the parents of both head and spouse have either died or do not reside in the same village.

of family networks. As reported in Table 4, on average, there are 7.8 households within the same family network. Table 2 previously documented a household is connected, at distance one, to 5.23 other households in the village, suggesting on average, each household has two to three ties of distance two or more in the village. This highlights that, as expected, extended family networks are dense in rural village economies in the sense that those any given household i is connected to, are likely to also be connected to each other. Relative to the size of the village, 17% of households within a given village are part of the same family network, and Figure 2C shows the overall variation in the share of village households that are part of the same dynasty.

Taken together this descriptive evidence highlights that in the long run, there remain distinct family networks within most villages. It is not therefore the case that through the marriage market, all households become connected to all others in physical proximity. This is the case despite the sampled village being of relatively small size.

Table 4 provides further information on statistics that describe the family network. The average distance between any two households in the same network is on average, 2.12. The degree of the average network, defined as the number of households each is directly connected to, is 2.07. This is smaller than the average number of links each household in the sample has, as reported in Table 3, because in Table 4 there is one observation per family network and all are equally weighted irrespective of the number of households that are embedded within the network. This highlights that there are many family networks that comprise only two households, as shown in Figures 3B and 4. The diameter of the network, defined as the largest distance d_{ij} between any two households i and j in the network, is around 2.5 on average. One implication of this is that family networks are unlikely to span across more than three generations.

The final column presents evidence on the welfare index of households within the same network. As this is standardized within state, and all households in the family network are defined to be in the same village and hence state, by construction this has mean zero. More important is the considerable variation in household welfare within the same extended family network – this variation is almost as great as that across villages. This highlights the need for, and availability of, insurance opportunities within the same family network. This in turn may drive the formation or sustainability of family networks in risky economic environments. In addition, this highlights that there may be assortative matching in marriage markets by family wealth. Hence there remains considerable variation in welfare levels across family networks within the same village.

The table also decomposes the variation in each statistic into that between villages and that within the same village. We note there is considerable variation arising from both dimensions of the data. Hence there are considerable differences in the structure of family networks in different villages, and in the network structure of family dynasties that reside in the same village.

To emphasize how the structure of family networks can differ between villages, Figure 4 provides a graphical representation of family networks in two villages of median size, corresponding to 36 households. The upper panel shows a village with a relatively disperse set of family networks. There are five families in which there are only two households present, one large family network in which there are ten households, and 16 households that have no family ties to any other household in the village. In contrast the lower panel shows another village – which recall is of the same size – but in which there exists one large family network containing 30 households, and 6 single node households. In this paper we focus on features of the village economy that relate to family networks in the village. We therefore shed light on what may be the underlying determinants of the types of network structure observed across villages, as highlighted in Figure 4.

4   Empirical Evidence

4.1   Extended Family Networks

We now turn to estimate the relationship between characteristics of the village economy on the overall number of family dynasties, the structure of these networks, and the specific types of family relations that reside within them. To begin with, we consider the effect of the village marginality index (M_v) and the village inequality index (Q_v), on the number of family networks in the village (N_v). We therefore estimate an OLS regression at the village level of the following form,

$$N_v = \beta M_v + \gamma Q_v + \delta S_v + u_v. \tag{4.1}$$

We also condition on the size of the village (S_v) as there is a somewhat mechanical relationship between village size and the number of family dynasties present within it. Throughout, robust standard errors are calculated. The results are reported in Table 5.

Column 1 shows that – (i) in villages that are more marginal – namely worse off – there are significantly fewer family dynasties present; (ii) in villages that have greater levels of inequality, there are significantly more family dynasties present; (iii) larger villages have more family dynasties resident within them as expected.

As villages in the sample as geographically dispersed over seven states and 115 municipalities of Southern Mexico, some of the variation in the economic environment across villages relates to differences across states and municipalities. To purge our estimates of these and to be able to explore the effects of the village economy on family networks within more concentrated geographic locations, the specifications in Columns 2 and 3 control for state and municipality fixed effects respectively. The previous results are robust to these specifications, so that within the same state

or municipality, more marginal villages, and those that are more equal in terms of household welfare, are associated with fewer family dynasties being resident in them.

To shed more light on this, Columns 4 and 5 focus on how these characteristics of the village economy relate to the size of the largest family dynasty, both in terms of the absolute number of households embedded within it, and the share of households in the village that are embedded within the dominant family. The results show that within the same municipality – (i) in more marginal villages, the size of the largest network is significantly higher, as is the share of households within it; (ii) in more equal villages, the size of the largest network is significantly higher, as is the share of households within it.

The remaining columns focus on statistics, discussed in Section 3.3, that describe the family network as a whole – the average distance between any two households in the same network, the degree of the network, and the diameter of the network. We therefore estimate an OLS regression at the network level of the following form,

$$F_{nvm} = \alpha_m + \beta M_{vm} + \gamma Q_{vm} + \delta S_{vm} + u_{nvm}, \tag{4.2}$$

where F_{nvm} refers to each statistic, F, describing family network n in village v in municipality m, and α_m denotes municipality fixed effects that capture differences across family networks due to the geographic location of the village. All other variables are as previously defined, except the subscripts now make explicit that each village v is located within municipality m.

Columns 6 to 8 show that – (i) in more marginal villages, the average density or diameter of the network do not change overall, although the degree of each household within the network significantly increases; (ii) in more unequal villages, family networks are significantly less dense as measured by the average distance between any two households, the overall number of links each household has, and the diameter of the network; (iii) each network statistic, F_{nvm}, is positively correlated with the size of the village as expected.

The results highlight the existence of an important interplay between the characteristics of the local village economy, and the extended family networks present in the village. The most intuitive way in which to benchmark the magnitude of the effects is in terms of the number of households in the largest network. Taking the estimates from our preferred specification in Column 4 and the descriptive evidence from Table 1 reveals that – (i) if the village marginality index were to increase by one standard deviation, the largest family network comprises 1.66 more households; (ii) a one standard deviation decrease in the village inequality index is associated with there being 2.10 more households in the largest family network; (iii) the magnitude of the effect of the marginality index (inequality index) is equivalent to there being 12.6 (10.0) more households in the village *per se*, relative to a mean village size of 44.7 households.

The relationships identified in Table 5 therefore emphasize that extended family network structures appear to be sensitive to levels of welfare within the village, as well as to the dispersion of welfare across households within the village. Of course, as these relations are identified from a cross section of villages, then absent any exogenous variation in the marginality of the village or inequality within it, it is impossible to identify whether network structures endogenously respond to these characteristics of the village economy, or whether the features of the local economy are themselves shaped by the nature of extended family networks within them.

In either case, the results show that understanding the causes and consequences of why villages differ in their levels of poverty, the vulnerability of the local population to aggregate income shocks, and their integration into local markets – as captured in the village marginality index, or understanding the causes and consequences of why villages differ in their inequality in terms of income, assets, or land holdings – as measured by the village inequality index, both require an understanding of how extended family networks also differ across villages.

4.2   Extended Family Links

In the previous subsection we established that fewer family dynasties exist in more marginal and more equal villages, and that each dynasty comprises more households in absolute and relative terms. This helps shed light on one dimension in which the villages depicted in Figure 4, differ so markedly from each other, despite being of the same size. We now provide evidence on how the precise inter and intra-generational family links that exist in the village, are related to underlying features of the village economy.

To do this we define L_{jvm} to be the share of households in village v in municipality m that have an extended family link of type-j. As made precise in Section 2, exploiting the patronymic naming convention allows us to define the following eight types of family link j – (i) intra-generational family links, such as those from the head (spouse) of the household to his (her) brothers and sisters; (ii) inter-generational family links, such as those from the head (spouse) of a household to his (her) parents, and from the head and spouse of the household to their adult sons and daughters. We then estimate the following specification,

$$L_{jvm} = \alpha_m + \beta M_{vm} + \gamma Q_{vm} + \delta S_{vm} + u_{jvm}, \tag{4.3}$$

where all other variables are as previously defined. An econometric concern with this specification is that the presence of a given link type-j is likely to be influenced by the presence of other link types, j', so that in effect we need to estimate a linear equation for each j. We therefore use seemingly unrelated regression (SUR) methods to estimate these j-equations allowing for cross-

equation correlation in the error terms [Zellner 1962]. The results are reported in Table 6.[12]

To begin with, we focus on the results in the first four columns on the share of households in the village that have inter-generational extended family links. We see that the proportion of households in the village economy that have either their parents present, or their adult sons and daughters within the village, is largely insensitive to characteristics of the village economy, as captured by the village marginality and inequality indices. As expected, the share of households with such links is significantly increasing in the size of village.

The second set of results on the presence of intra-generational links, namely siblings, show more robust correlations with features of the village economy. In particular, for three out of four sibling link types, more marginal villages have significantly more households present in them that have at least one tie to a sibling. For all sibling types, more equal villages are associated with households being significantly more likely to have at least one sibling present in the village.

Taken together with the results in Table 4, the evidence points to the reason behind why family networks are larger in more marginal and more equal villages, being that adult siblings are more likely to reside within the same village, all else equal.

To complement these results, the last panel of Table 6 uses information from the household roster to identify extended family members that co-reside in the household, and then estimates a specification analogous to (4.3) for the share of households in the village in which at least one type-j link resides inside the household. As discussed in Section 3, it is predominantly inter-generational links that reside within the household so we estimate a SUR specification for these four types of link. By combining the results in Columns 1 to 4 on inter-generational links outside of the household and within the village, with these results in Columns 9 to 12 on such links coresident in the household, helps shed light on the nature of household partition [Foster 1993].

On parental links, the results show that – (i) in more marginal villages, the parents of the head are slightly less likely to reside within another household in the same village and are slightly more likely to remain within the household (Columns 1 and 9); (ii) the presence of the spouse's parents outside the household is unrelated to features of the village economy (Column 2) although they are more likely to reside within the household in more unequal villages (Column 10).

On links to adult children, we see that the presence of adult daughters within the household is particularly sensitive to characteristics of the village economy. In more marginal and more equal villages, adult daughters are significantly less likely to reside within the household (Column 12),

---

[12]A concern that arises with this specification is that the dependent variable is a share and so lies between zero and one, which is not accounted for by the SUR estimates. In relation to this we note that the mean of the dependent variable lies well within the zero-one interval, as shown in Table 2. Moreover, we note that few predicted values lie outside of the unit interval. For each equation, less than 2% of the village observations have a predicted value less than zero, and none have a predicted value greater than one.

and as shown earlier, more likely to reside in another household within the same village (Column 4). This is consistent with daughters being less likely to marry and move outside of the village, in local economies characterized by more extreme poverty and inequality across households. In contrast, the likelihood an adult son is present – either within the household or in another household within the same village – is less sensitive to features of the village economy (Columns 3 and 11).[13]

4.3   Household Level Analysis

One interpretation of the results is that households sort into villages on the basis of the characteristics of the village economy. Hence the relationships we find between the village marginality and village inequality indices and the structure of extended family networks, may reflect that different types of household reside in villages that are less integrated into the market economy or more unequal, for example. To explore this issue we now move to a household level analysis to shed light on whether, conditional on household characteristics, features of the village economy continue to have robust correlations with the presence of extended family members. We therefore estimate the following specification for whether household h in village v in municipality m has at least one type-j link to another household in the same village,

$$L_{jhvm} = \alpha_m + \beta M_{vm} + \gamma Q_{vm} + \delta S_{vm} + \lambda X_{hvm} + u_{jhvm}, \tag{4.4}$$

where X_{hvm} are household characteristics, and all other controls are as previously defined. As before, we estimate this system of eight equations for each inter and intra-generational link defined within a SUR framework.

As discussed in the introduction, a large literature has explored the mechanisms driving the presence of extended family members. Hence it is important to condition on correlates of such factors in X_{hvm}. This allows us to then explore whether features of the village economy are related to the presence of extended family members over and above such household observables.

We condition on the following controls in X_{hvm}, as defined in the first wave of household data collected in October 1997. First, there is a somewhat mechanical correlation between the age of the head and spouse and the likelihood that their parents and adult children are in close proximity

---

[13]Table A1 reports the correlation matrix of residuals from these SUR estimates of links outside and inside the household. The Breusch-Pagan test of independence of the residuals for extended family links outside the household has a p-value of .000. This is also the case for inter-generational links inside the household. The correlation matrix shows that for all pairs of links outside the household the correlation in residuals is positive. The correlation among inter-generational (intra-generational) links is at least .233 (.364). For inter-generational links within the household, the corresponding correlations are smaller in magnitude and sometimes negative, as shown in parentheses. The highest correlation for such links is between the residuals for the presence of adult sons and daughters in the household.

so we control for the ages of the head and spouse. To capture any ethnic differences across families, we control for whether the head and spouse speak an indigenous language. We also condition on the number of members resident in the household.

Second, wealthier family dynasties may enjoy higher fertility and lower mortality, and so are more likely to have extended family members present, other things equal. Wealthier families are more likely to own land and if land markets are missing, the ability to inherit land, or to acquire land specific human capital, may lead adult children to be more likely to remain within the village than otherwise [Foster and Rosenzweig 2002]. We therefore control for the years of completed schooling of the head and spouse, whether the head and spouse are working, whether the household owns its own home, any land, any livestock, whether the house has floors of dirt, and the household welfare index.[14]

A third mechanism behind why households may partition and form larger extended family networks is through the need to insure against idiosyncratic income shocks. Such insurance can be gained through the strategic marriage of daughters into families with less correlated income shocks, or through the migration of some family members to other locations [Rosenzweig and Stark 1989, Foster 1993]. We therefore control for whether any members of the household have permanently migrated away in the five years prior to October 1997, and whether any household members have temporarily migrated away in the twelve months prior to October 1997.

Table 7 reports the results. We see that, controlling for a rich set of household characteristics, it continues to be the case that features of the village economy – as captured in the village marginality and inequality indices – are significantly correlated to the presence of extended family members both outside the household and within the village (Columns 1 to 8), and coresident within the household (Columns 9 to 12). For each family link type-j, there are significantly more such links in larger villages as expected. Moreover, the estimated effects of these features of the local economy match quite closely with those obtained earlier at the village level in Table 6.

More precisely, on ties to parents, the presence of the parents of the head and spouse are not much influenced by features of the village economy. This applies equally to whether they are resident within the village (Columns 1 and 2) or coresident in the household (Columns 9 and 10).

On ties to adult sons, in more marginal or more equal villages, adult sons are significantly more likely to reside both within the village or be coresident within the household (Columns 3 and 11). Hence this suggests adult sons are less likely to move away from villages that are more marginal or more equal.

The results are slightly different for links to adult daughters. While it remains the case, as for

---

[14]We do not control for literacy because it is highly correlated with years of schooling – 89% (90%) of heads (spouses) have no formal schooling if they are illiterate.

adult sons, that in more marginal or more equal villages, adult daughters are significantly more likely to reside within the village (Column 4), unlike sons, these features lead them to be less likely to remain coresident within the household (Column 12). This is consistent with the idea that migrants do not originate from the poorest households.

These results suggest the partition of households, as driven by the movement of adult daughters, is such that they are less likely to marry outside of the village, in local economies that are more marginal or more equal. This raises the possibility that marriages of daughters are used as strategic devices to geographically diversify risk. In contrast, the location choices of males may result more from the non-tradeability of labor exchange across villages, and the nature of household production that is tied to specific plots of land within the village.

On ties to siblings, we see that in line with the village level analysis in Table 6, adult siblings of both the head and spouse are significantly more likely to reside within the same village in local economies that are more marginal or less unequal (Columns 5 to 8).[15]

Taken together, these results highlight that the structure of family networks relates to features of the village economy as a whole, over and above the characteristics of individual households within them.[16]

4.4 Other Features of the Village Economy

Throughout we have focused on the relationship between extended family structures and two features of the local economic environment – (i) the marginality of the village, a summary statistic characterizing average levels of poverty within the village, the vulnerability of the local population to aggregate income shocks, and the integration of the village economy into local markets and

---

[15]Two further points are of note. First, the set of households in Table 7 include both connected and isolated households. The results are robust to restricting the sample to only those 78% of households that are embedded within a family network. Second, although not the focus of this paper, we note that the correlations between household characteristics and the presence of family ties have intuitive signs. For example, the mechanical correlations with age are as expected with older heads and spouses being less likely to have their parents outside of the household and resident in the village, and significantly more likely to have their adult children in other households in the village. Heads and spouses with more years of schooling are more likely to have their parents present. Households from which there are seasonal migrants are also more likely to have inter-genarational links present in the village, although there is no effect on the presence of siblings. Finally, whether the head and spouse speak an indigenous language does not predict the presence of extended family ties. This is reassuring because the number of extended family ties, for each type of tie, are no different between indigenous and non-indigenous households.

[16]As a check on these results we also explored whether the features of the village economy have heterogeneous effects across households on the formation of family networks. To do so we use the classification of households into poor and non-poor designed to determine eligibility for *Progresa* transfers. This classification is made in 1997 on the basis of the household poverty index. We then estimate (4.4) allowing the effects of M_{vm} and Q_{vm} to differ between those 52% of households classified as poor, and the remaining 48% classified as non-poor. We find no robust evidence that, conditional on household characteristics, the features of the aggregate village economy have heterogeneous effects across poor and non-poor households.

the wider economy; (ii) village inequality, a summary statistic characterizing inequality across households in income, acquired human capital, assets, and land holdings. However, there are potentially many other features of the local economy that relate to the formation and structure of family networks.

One aspect of the economic environment that is not well captured in either index is the level of aggregate uncertainty faced by households in the village. In order to explore this issue we exploit information from the community surveys on the prevalence of natural shocks that afflict the village. In the last three surveys – October 1998, March 1999, November 1999 – information was collected on whether each village was affected by the following types of shock – droughts, floods, frosts, fires, pests, earthquakes, or hurricanes.[17]

We use this information to calculate the average number of such shocks the village has been exposed to over each 6 month period. This measure then serves as a proxy of aggregate uncertainty, or the persistence of shocks, at the village level, assuming the prevalence of such shocks from October 1998 to November 1999 reflects their long run incidence. In our sample of villages, the number of biannual shocks varies from zero to 3.67. The average (median) village is affected by 1.30 (1.33) shocks over a 6 month period, with standard deviation 6.87. We then re-estimate (4.4) and additionally control for this proxy of aggregate uncertainty at the village level.[18]

Table 8 presents the results. Given the addition control relates to exogenously occurring natural shocks, the coefficients on the village marginality and village inequality indices are both unaffected by additionally controlling for this measure of aggregate uncertainty. While there is no evidence that the incidence of natural shocks is related to the presence of parental links outside of the household, they are more robustly related to the presence of sibling ties within the village. This is especially the case for all sibling links that involve men. Combining these results with those in Columns 9 to 12 on coresidence within the household, we see that in villages that have a higher prevalence of naturally occurring shocks, adult sons are less likely to remain within the household and more likely to form a new household within the village (Columns 3 and 11). In contrast, the presence of adult daughters – either within the household or within the village – is unaffected by these natural shocks (Columns 4 and 12).[19]

---

[17]We focus on naturally occurring shocks as the incidence of these is unaffected by any endogenous change in family network structures due to *Progresa*, or directly because of the program.

[18]Using the average number of shocks per six month period produces identical qualitative results to using the total number of shocks over the entire 18 month period as they are linear transformations of each other, and each village is surveyed three times.

[19]To shed light on the qualitative importance of aggregate uncertainty – as captured in the incidence of natural shocks – vis-à-vis the other measures of the economic environment we have used throughout, we focus attention on the effects on the presence of adult sons (Column 3). The results imply the effects of a one standard deviation increase in the village marginality or village inequality indices, both change the likelihood of at least one adult son being present in the village, by an amount that is approximately equivalent to there being one more natural shock

5   Conclusions

We have presented evidence on the empirical correlations between features of the economic environment in rural villages, and the extended family network structures that are present there. Using household and community data from 500 villages in rural Mexico, we have documented the relationship between family networks and two features of the village economy. The first is an index of the marginality of the village – a summary statistic that captures average levels of poverty within the village, the vulnerability of the local population to aggregate income shocks, and the integration of the village economy into local markets and the wider economy. The second is an index of village level inequality – a summary statistic characterizing inequality across households in income, acquired human capital, assets, and land holdings.

We find that in more marginal villages, and villages in which households are equally poor, there are significantly fewer family dynasties present. Each dynasty is of larger size both in terms of the absolute number of households, and the share of village households that are embedded within the network. The reason why family networks are larger in such villages is that heads of household and their spouses are both significantly more likely to have their siblings and adult sons and daughters resident within the village. Finally, we find that controlling for a rich set of household characteristics, these features of the village economy continue to relate to family network structures over and above the characteristics of individual households within them.

The results emphasize the considerable heterogeneity *across* villages both in terms of their economic characteristics, but also in terms of the extended family structures that reside within them. Hence for example, finding heterogeneous effects of large scale policy interventions in rural economies, may be as much due to across village differences in network structures, as due to variation in more easily observed dimensions of the local economy. The results also emphasize the considerable heterogeneity *within* villages in terms of whether households are embedded into extended family networks or isolated in the sense that none of their extended family are geographically proximate. Moreover, among connected households, there remains considerable variation in the overall structure of their network, as well as in their specific inter and inter-generational links.

Taken together, one interpretation of this evidence is that in economic environments characterized by – (i) imperfect markets, correlated income shocks, an absence of formal institutions of contract enforcement; and, (ii) where all households are equally poor, one strategy of the poor is to surround themselves with extended family members. The benefits of having a geographically dispersed network – such as being subject to less correlated income shocks due to climatic conditions – seem to be outweighed, on average, by the costs of such geographic dispersion – such

---

to the village every six months.

as increased transactions costs of monitoring and making non-monetary transfers to others in the network.

A natural question that arises from this analysis is whether, ultimately, the presence of extended family members influences household behavior and outcomes. An existing literature provides evidence on the effect of family links on household behavior, such as for consumption [Altonji *et al* 1992], inter-generational transfers [Cox and Jakubson 1995, Altonji *et al* 1997, La Ferrara 2003, Behrman and Rosenzweig 2006], children's education choices [Loury 2006], non-resident parental investments into children [Weiss and Willis 1985], and the crowding out of private transfers by public transfers [Albarran and Attanasio 2003, Cox and Fafchamps 2007]. Relatedly, Strauss and Thomas [1995] review the empirical evidence on whether mothers behave more altruistically towards their children than fathers, and Duflo [2003] presents evidence consistent with maternal and paternal grandmothers having different degrees of altruism towards their grandchildren.

In relation to the empirical setting in this paper, in Angelucci *et al* [2007] we identify whether the presence and characteristics of extended family members influence household responses, in terms of secondary school enrolment, to the *Progresa* program. One important component of the program is that it provides cash transfers to households conditional on the secondary enrolment of their children. We document that only eligible households that are embedded within family networks respond to the program along this margin. In contrast, households who are eligible but isolated – in the sense that none of their family members live in the same village – do not respond along this margin, and that the difference in the behavioral responses of connected and isolated households within the same village is qualitatively significant and quantitatively large.

In terms of a broader research agenda, if extended families shape the objectives and constraints relevant for households within them, then neglecting the role of the extended family when analyzing household behavior may lead to an incomplete understanding of the forces driving household choices and outcomes. The extended family may be especially relevant in economic environments characterized by missing markets, correlated income shocks, and the prevalence of informal institutions of contract enforcement. The evidence in this paper suggests there is an important interplay between such features of the village economy, and the structure of family networks that reside there. An important and challenging next step is to find and exploit exogenous sources of variation in either dimension, to more precisely identify the causal relationship between the two.

References

[1] ALBARRAN.P AND O.ATTANASIO (2003) "Limited Commitment and Crowding Out of Private Transfers: Evidence from a Randomised Experiment", *Economic Journal* 113: C77-85.

[2] ALTONJI.J.G, F.HAYASHI, AND L.J.KOTLIKOFF (1992) "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data", *American Economic Review* 85: 1177-98.

[3] ALTONJI.J.G, F.HAYASHI, AND L.J.KOTLIKOFF (1997) "Parental Altruism and Inter Vivos Transfers: Theory and Evidence", *Journal of Political Economy* 105: 1121-66.

[4] ANGELUCCI.M, G.DE GIORGI, M.A.RANGEL, AND I.RASUL (2007) Family Networks and School Enrolment: Evidence from a Randomized Social Experiment, mimeo University College London.

[5] BANERJEE.A, D.MOOKHERJEE, K.MUNSHI, AND D.RAY (2001) "Inequality, Control Rights and Efficiency: A Study of Sugar Cooperatives in Western Maharashtra", *Journal of Political Economy* 109: 138-90.

[6] BARDHAN.P (1984) *Land, Labor, and Rural Poverty*, New York: Columbia University Press.

[7] BARDHAN.P, M.GHATAK AND A.KARAIVANOV (2002) "Inequality, Market Imperfections, and the Voluntary Provision of Collective Goods", in Baland.J, P.Bardhan and S.Bowles (eds.), *Inequality, Collective Action and Environmental Sustainability*, Russell Sage Foundation and Princeton University Press.

[8] BASU.K (1986) "One Kind of Power", *Oxford Economic Papers* 38: 259-82.

[9] BEHRMAN.J.R AND M.ROSENZWEIG (2006) "Parental Wealth and Adult Children's Welfare in Marriage", *Review of Economics and Statistics* 88: 496-509.

[10] COX.D AND M.FAFCHAMPS (2007) "Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions", forthcoming *Handbook of Development Economics* Volume 4.

[11] COX.D AND G.JAKUBSON (1995) "The Connection Between Public Transfers and Private Interfamily Transfers", *Journal of Public Economics* 57: 129-67.

[12] DEATON.A (1992) "Saving and Income Smoothing in Côte D'Ivoire", *Journal of African Economies* 1: 1-24.

[13] DERCON.S AND P.KRISHNAN (2000) "In Sickness and in Health: Risk Sharing Within Households in Rural Ethiopia", *Journal of Political Economy* 108: 688-727.

[14] DUFLO.E (2003) "Grandmothers and Granddaughters: Old Age Pension and Intra-household Allocation in South Africa", *World Bank Economic Review* 17: 1-25.

[15] DUFLO.E AND A.BANERJEE (2007) "The Economic Lives of the Poor", *Journal of Economic Perspectives* 21: 141-67.

[16] ELLSWORTH.L (1989) Mutual Insurance and Non-Market Transactions Among Farmers in Burkina Faso", University of Wisconsin, Ph.D. Dissertation.

[17] ERMISCH.J.F, M.FRANCESCONI, AND T.SIEDLER (2006) "Intergenerational Mobility and Marital Sorting", *Economic Journal* 116: 659-79.

[18] FAFCHAMPS.M AND S.LUND (2003) "Risk-sharing Networks in Rural Philippines", *Journal of Development Economics* 71: 261-87.

[19] FERNANDEZ.R, N.GUNER, AND J.KOWLES (2005) "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality", *Quarterly Journal of Economics* 120: 273-344.

[20] FOSTER.A (1993) "Household Partition in Rural Bangladesh", *Population Studies* 47: 97-114.

[21] FOSTER.A AND M.ROSENZWEIG (2002) "Household Division and Rural Economic Growth", *Review of Economic Studies* 69: 839-69.

[22] KREMER.M (1997) "How Much Does Sorting Increase Inequality?", *Quarterly Journal of Economics* 112: 115-39.

[23] KRISHNAN.P AND E.SCIUBBA (2004) Endogenous Network Formation and Informal Institutions in Village Economies, Cambridge Working Papers in Economics, 462.

[24] LA FERRARA.E (2003) "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana", *American Economic Review* 93: 1730-51.

[25] LIGON.E (1998) "Risk-Sharing and Information in Village Economies", *Review of Economic Studies* 65: 847-64.

[26] LOURY.L.D (2006) "All in the Extended family: Effects of Grandparents, Aunts, and Uncles on Educational Attainment", *American Economic Review Papers and Proceedings* 96: 275-8.

[27] LUCAS.R AND O.STARK (1985) "Motivations to Remit: Evidence from Botswana", *Journal of Political Economy* 93: 901-18.

[28] MUNSHI.K AND ROSENZWEIG.M (2005) Why is Mobility in India so Low? Social Insurance, Inequality, and Growth, mimeo Brown University.

[29] OLSON.M (1965) *The Logic of Collective Action*, Harvard University Press, Cambridge: Massachusetts.

[30] ROSENZWEIG.M.R (1988) "Risk, Implicit Contracts and the Family in Rural Areas of Low Income Countries", *Economic Journal* 98: 1148-70.

[31] ROSENZWEIG.M AND O.STARK (1989) "Consumption Smoothing, Migration, and Marriage: Evidence From Rural India", *Journal of Political Economy* 97: 905-26.

[32] STRAUSS.J AND D.THOMAS (1995) "Human Resources: Empirical Modelling of Household and Family Decisions", in J.R.Behrman and T.N.Srinivasan (editors) *Handbook of Development Economics* Volume 3, Amsterdam: Elsevier Science.

[33] TOWNSEND.R (1994) "Risk and Insurance in Village India", *Econometrica* 62: 539-91.

[34] UDRY.C (1994) 'Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria', *Review of Economic Studies* 61: 495-526.

[35] WASSERMAN.S AND K.FAUST (1994) *Social Network Analysis: Methods and Applications*, Cambridge: Cambridge University Press.

[36] WEISS.Y AND R.J.WILLIS (1985) "Children as Collective Goods and Divorce Settlements", *Journal of Labor Economics* 3: 268-92.

[37] ZELLNER.A (1962) "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias", *Journal of the American Statistical Association* 57: 348-68.

Table 1: The Village Economy

Mean, standard deviation in parentheses

	Village marginality index	SD of household welfare index in the village (Village inequality index)	Village size (number of households)
Village Characteristic	-.003 (.994)	.840 (.190)	44.7 (29.2)

Notes: There is one observation for each village. The village marginality index is constructed from information on the share of illiterate adults in the village, the share of dwellings without water, drainage systems, electricity, and with floors of dirt, the average number of occupants per room in village households, the share of the population working in the primary sector, distances from other villages, and health and school infrastructures located in the village. A higher marginality index corresponds to the village being more marginal (poorer). The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. An increase in the index implies the households is less poor. The measure of village inequality is the standard deviation of the welfare index of all households in the village. Village size is defined as the number of households in the village. The village marginality index is standardized across all villages. The household welfare index is defined relative to a state norm. Hence the village inequality index is standardized at the state level. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village. The villages in the sample cover 7 states and 115 municipalities.

Table 2: The Share of Households With Extended Family Links in the Village, by Type of Link

Mean, standard deviation in parentheses

	Any Family Link (Connected)	Inter-generational Family Links				Intra-generational Family Links			
		Parents to Son	Parents to Daughter	Son to Parent	Daughter to Parent	Head to Head (Brothers)	Head to Spouse	Spouse to Head	Spouse to Spouse (Sisters)
Fraction of households in the village with such a family link	.781	.164	.085	.178	.087	.452	.285	.262	.230
	(.156)	(.100)	(.067)	(.108)	(.068)	(.186)	(.173)	(.150)	(.164)

Notes: There is one observation for each village. The sample is based on couple headed households that can be tracked over the first and third Progresa waves. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village.

Table 3: The Number of Extended Family Links of Households, by Type of Link

Connected Households

Means, standard deviation between villages in parentheses, standard deviation within villages in brackets

	Parents	Adult Children		Siblings		All Extended Family Links	Total Links From Household
	Father and Mother	Sons	Daughters	Brothers	Sisters		
From head of household to:	.461			1.39	.840	3.34	
	(.094)	.443	.209	(.340)	(.325)	(1.02)	5.23
	[.837]	(.399)	(1.95)	[1.87]	[1.52]	[3.92]	(1.53)
From spouse of household to:	.250	[1.44]	[.813]	.910	.725	2.54	[5.47]
	(.260)			(1.23)	(.304)	(1.04)	
	[.656]			[1.71]	[1.51]	[3.89]	

Notes: There is one observation per household. The sample is restricted to couple headed households that can be tracked over the first and third Progresa waves that have at least one extended family member present in the village. An adult child defined is to be at least 17 years old. By construction, the number of links to adult sons and daughters is the same from the head and spouse. The total links from the household is the sum of unique extended family links from the household to others in the same village, and does not therefore double count the children of the head and spouse of the household. The decomposition of the standard deviation into that between and within villages takes account of the fact that the number of households differs in each village.

Table 4: Extended Family Network Descriptives

Means, standard deviation between villages in parentheses, standard deviation within villages in brackets

	Network Size	Network Size/Number of Households in Village	Average Distance	Degree	Diameter	Standardized Household Welfare Index
Mean	7.76	.167	2.12	2.07	2.45	-.001
SD between villages	(9.65)	(.149)	(.303)	(1.20)	(1.18)	(.566)
SD within villages	[11.3]	[.153]	[.545]	[1.19]	[2.11]	[.521]

Notes: There is one observation per family network so that each network has the same weight irrespective of the number of households within it. The underlying sample of households is based on couple headed households that can be tracked over the first and third Progresa waves. Of the baseline sample of 22553 couple headed households, 17030 (75.5%) of them are within family networks with at least two households. There are 2196 family networks in total. The size of the network is the number of households in the network. Two households that are directly connected are defined to be of distance one to each other. The average distance is the average over all possible pairs of households within the family network. The diameter of the networks is the longest distance between two households that exists in the network. The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. The index is calculated relative to a state norm. Hence we standardize the index within each state. An increase in the index implies the households is less poor. The decomposition of the standard deviation into that between and within villages takes account of the fact that the number of family networks differs in each village.

Table 5: Family Network Structures and Village Characteristics

OLS regression estimates, robust standard errors in parentheses

Dependent Variable:	Village Level Statistics					Network Level Statistics		
	Number of Family Networks			Size of the Largest Family Network	Share of Village Households that are in the Largest Family Network	Average Distance	Degree	Diameter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Village marginality index	-.237*	-.253**	-.406**	1.67**	.029**	-.015	.217***	.038
	(.129)	(.124)	(.164)	(.809)	(.014)	(.019)	(.059)	(.073)
Village inequality index	2.27***	2.01***	1.79**	-11.1***	-.219***	-.194**	-.700**	-.684**
	(.559)	(.582)	(.718)	(3.73)	(.071)	(.086)	(.276)	(.347)
Village size	.339***	.346***	.347***	.718***	.230***	.003***	.005**	.010***
	(.721)	(.736)	(.770)	(5.40)	(.045)	(.001)	(.002)	(.003)
State fixed effects	No	Yes	No	No	No	No	No	No
Municipality fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.167	.197	.243	.749	.325	.026	.098	.019
Observations	504	504	504	504	504	2196	2196	2196

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. There is one observation for each village, and each dependent variable is constructed using couple headed households that can be tracked over the first and third Progresa waves in each village. The village marginality index is constructed from information on the share of illiterate adults in the village, the share of dwellings without water, drainage systems, electricity, and with floors of dirt, the average number of occupants per room in village households, the share of the population working in the primary sector, distances from other villages, and health and school infrastructures located in the village. A higher marginality index corresponds to the village being more marginal (poorer). The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. An increase in the index implies the households is less poor. The measure of village inequality is the standard deviation of the welfare index of all households in the village. Village size is defined as the number of households in the village. The village marginality index is standardized across all villages. The household welfare index is defined relative to a state norm. Hence the village inequality index is standardized at the state level. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village. The villages in the sample cover 7 states and 115 municipalities. Of the baseline sample of 22553 couple headed households, 17030 (75.5%) of them are within family networks with at least two households. There are 2196 family networks in total. The size of the network is the number of households in the network. Two households that are directly connected are defined to be of distance one to each other. The average distance is the average over all possible pairs of households within the family network. The degree is the number of households the average household in the family network is connected to. The diameter of the networks is the longest distance between two households that exists in the network. Robust standard errors are reported.

Table 6: Extended Family Links and Village Characteristics

Seemingly Unrelated Regression Estimates

	Outside the Household								Inside the Household			
	Inter-generational Family Links				Intra-generational Family Links				Inter-generational Family Links			
Dependent Variable:	Son to Parent	Daughter to Parent	Parents to Son	Parents to Daughter	Head to Head (Brothers)	Head to Spouse	Spouse to Head	Spouse to Spouse (Sisters)	Son to Parent	Daughter to Parent	Parents to Son	Parents to Daughter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Village marginality index	-.010*	.002	.007	.007**	.008	.018**	.028***	.026***	.006*	-.001	-.003	-.032***
	(.005)	(.004)	(.005)	(.003)	(.009)	(.008)	(.007)	(.007)	(.003)	(.001)	(.006)	(.006)
Village inequality index	-.037	-.012	-.061**	-.017	-.152***	-.146***	-.082**	-.139***	-.005	.021***	.045	.086***
	(.026)	(.017)	(.025)	(.017)	(.044)	(.038)	(.033)	(.035)	(.016)	(.002)	(.028)	(.028)
Village size/100	.029*	.060***	.068***	.066***	.143***	.228***	.206***	.198***	-.004	.002	-.045***	-.046***
	(.015)	(.010)	(.014)	(.010)	(.025)	(.022)	(.019)	(.020)	(.009)	(.004)	(.016)	(.016)
Municipality fixed effects					Yes					Yes		
Observations					504					504		

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. There is one observation for each village, and each dependent variable is constructed using couple headed households in each village. The village marginality index is constructed from information on the share of illiterate adults in the village, the share of dwellings without water, drainage systems, electricity, and with floors of dirt, the average number of occupants per room in village households, the share of the population working in the primary sector, distances from other villages, and health and school infrastructures located in the village. A higher marginality index corresponds to the village being more marginal (poorer). The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. An increase in the index implies the households is less poor. The measure of village inequality is the standard deviation of the welfare index of all households in the village. Village size is defined as the number of households in the village. The village marginality index is standardized across all villages. The household welfare index is defined relative to a state norm. Hence the village inequality index is standardized at the state level. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village. The villages in the sample cover 7 states and 115 municipalities. In both SUR specifications, the Breusch-Pagan test of independence rejects the null hypothesis that the error terms are uncorrelated in each regression at the 1% significance level.

Table 7: Extended Family Links and Village and Household Characteristics

Seemingly Unrelated Regression Estimates

	Outside the Household								Inside the Household			
	Inter-generational Family Links				Intra-generational Family Links				Inter-generational Family Links			
Dependent Variable:	Son to Parent	Daughter to Parent	Parents to Son	Parents to Daughter	Head to Head (Brothers)	Head to Spouse	Spouse to Head	Spouse to Spouse (Sisters)	Son to Parent	Daughter to Parent	Parents to Son	Parents to Daughter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Village marginality index	.002	.003	.015***	.013***	.014**	.020***	.016***	.016***	.002	-.002	.022***	-.015***
	(.004)	(.004)	(.004)	(.003)	(.006)	(.006)	(.006)	(.005)	(.003)	(.002)	(.005)	(.005)
Village inequality index	-.033	-.010	-.109***	-.041**	-.191***	-.155***	-.099***	-.135***	-.002	.019**	-.085***	.029
	(.021)	(.017)	(.020)	(.016)	(.029)	(.027)	(.026)	(.025)	(.013)	(.008)	(.022)	(.022)
Village size/100	.015*	.037***	.055***	.051***	.108***	.184***	.163***	.151***	.002	-.004	-.011	-.018**
	(.008)	(.007)	(.008)	(.006)	(.011)	(.011)	(.010)	(.010)	(.005)	(.003)	(.009)	(.009)
Household controls					Yes						Yes	
Municipality fixed effects					Yes						Yes	
Observations					18979						18979	

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. There is one observation for each village, and each dependent variable is constructed using couple headed households in each village. The village marginality index is constructed from information on the share of illiterate adults in the village, the share of dwellings without water, drainage systems, electricity, and with floors of dirt, the average number of occupants per room in village households, the share of the population working in the primary sector, distances from other villages, and health and school infrastructures located in the village. A higher marginality index corresponds to the village being more marginal (poorer). The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. An increase in the index implies the households is less poor. The measure of village inequality is the standard deviation of the welfare index of all households in the village. Village size is defined as the number of households in the village. The village marginality index is standardized across all villages. The household welfare index is defined relative to a state norm. Hence the village inequality index is standardized at the state level. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village. The villages in the sample cover 115 municipalities. The following household level characteristics are also controlled for - the husband's age, years of schooling, whether he speaks an indigenous language, and whether has is currently working, the spouse's age, years of schooling, whether she speaks an indigenous language, and whether she is currently working, the household welfare index, the number of members of the household, whether the household owns its home, own land, has dirt floors, has any livestock, whether any member has permanently migrated away in the last five years, and whether any members have temporarily migrated away in the past year. In both SUR specifications, the Breusch-Pagan test of independence rejects the null hypothesis that the error terms are uncorrelated in each regression at the 1% significance level.

Table 8: Extended Family Links and the Prevalence of Natural Shocks

Seemingly Unrelated Regression Estimates

	Outside the Household								Inside the Household			
	Inter-generational Family Links				Intra-generational Family Links				Inter-generational Family Links			
Dependent Variable:	Son to Parent	Daughter to Parent	Parents to Son	Parents to Daughter	Head to Head (Brothers)	Head to Spouse	Spouse to Head	Spouse to Spouse (Sisters)	Son to Parent	Daughter to Parent	Parents to Son	Parents to Daughter
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Village marginality index	.001	.003	.016***	.013***	.015**	.021***	.017***	.017***	.003	-.002	.022***	-.015***
	(.004)	(.004)	(.004)	(.003)	(.006)	(.006)	(.006)	(.005)	(.003)	(.002)	(.005)	(.005)
Village inequality index	-.034	-.012	-.104***	-.040**	-.192***	-.155***	-.100***	-.139***	-.003	.019**	-.085***	.029
	(.022)	(.017)	(.020)	(.016)	(.029)	(.027)	(.026)	(.025)	(.013)	(.008)	(.022)	(.022)
Prevalence of natural shocks	.001	-.003	.015***	.005	.015**	.020***	.014**	.007	-.001	.004**	-.011**	-.003
	(.005)	(.004)	(.005)	(.004)	(.007)	(.006)	(.006)	(.006)	(.003)	(.002)	(.005)	(.005)
Village size/100	.018**	.039***	.051***	.051***	.103***	.184***	.163***	.149***	.003	-.004	-.012	-.018**
	(.009)	(.007)	(.008)	(.006)	(.011)	(.011)	(.010)	(.010)	(.005)	(.003)	(.009)	(.009)
Household controls					Yes						Yes	
Municipality fixed effects					Yes						Yes	
Observations					18979						18979	

Notes: *** denotes significance at 1%, ** at 5%, and * at 10%. There is one observation for each village, and each dependent variable is constructed using couple headed households in each village. The village marginality index is constructed from information on the share of illiterate adults in the village, the share of dwellings without water, drainage systems, electricity, and with floors of dirt, the average number of occupants per room in village households, the share of the population working in the primary sector, distances from other villages, and health and school infrastructures located in the village. A higher marginality index corresponds to the village being more marginal (poorer). The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. An increase in the index implies the households is less poor. The measure of village inequality is the standard deviation of the welfare index of all households in the village. Village size is defined as the number of households in the village. The village marginality index is standardized across all villages. The household welfare index is defined relative to a state norm. Hence the village inequality index is standardized at the state level. The measure of natural shocks is the average number of shocks in the village every six months (October 1998, March 1999, November 1999). These shocks include being affected by droughts, floods, frosts, fires, pests, earthquakes, or hurricanes. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village. The villages in the sample cover 115 municipalities. The following household level characteristics are also controlled for - the husband's age, years of schooling, whether he speaks an indigenous language, and whether has is currently working, the spouse's age, years of schooling, whether she speaks an indigenous language, and whether she is currently working, the household welfare index, the number of members of the household, whether the household owns its home, own land, has dirt floors, has any livestock, whether any member has permanently migrated away in the last five years, and whether any members have temporarily migrated away in the past year. In both SUR specifications, the Breusch-Pagan test of independence rejects the null hypothesis that the error terms are uncorrelated in each regression at the 1% significance level.

Table A1: Seemingly Unrelated Regression Estimate - Correlation Matrix of Residuals

|	Outside (Inside) the Household							
	Son to Parent	**Daughter to Parent**	**Parents to Son**	**Parents to Daughter**	**Head to Head (Brothers)**	**Head to Spouse**	**Spouse to Head**	**Spouse to Spouse (Sisters)**
Son to Parent	1							
Daughter to Parent	.327 (.135)	1						
Parents to Son	.492 (-.069)	.233 (-.024)	1					
Parents to Daughter	.248 (-.100)	.729 (-.000)	.425 (.390)	1				
Head to Head (Brothers)	.406	.231	.408	.311	1			
Head to Spouse	.271	.396	.408	.508	.541	1		
Spouse to Head	.159	.459	.405	.531	.364	.795	1	
Spouse to Spouse (Sisters)	.121	.315	.372	.342	.385	.486	.653	1

Notes: This correlation matrix of residuals corresponds to the SUR estimates in Table 6 of each extended family link regressed against the village marginality index, village inequality index, village size, and municipality fixed effects. The Breusch-Pagan test of independence of the residuals for extended family links outside the household has a p-value of .000. For links inside the household, the p-value is also .000.

Figure 1: The Village Economy
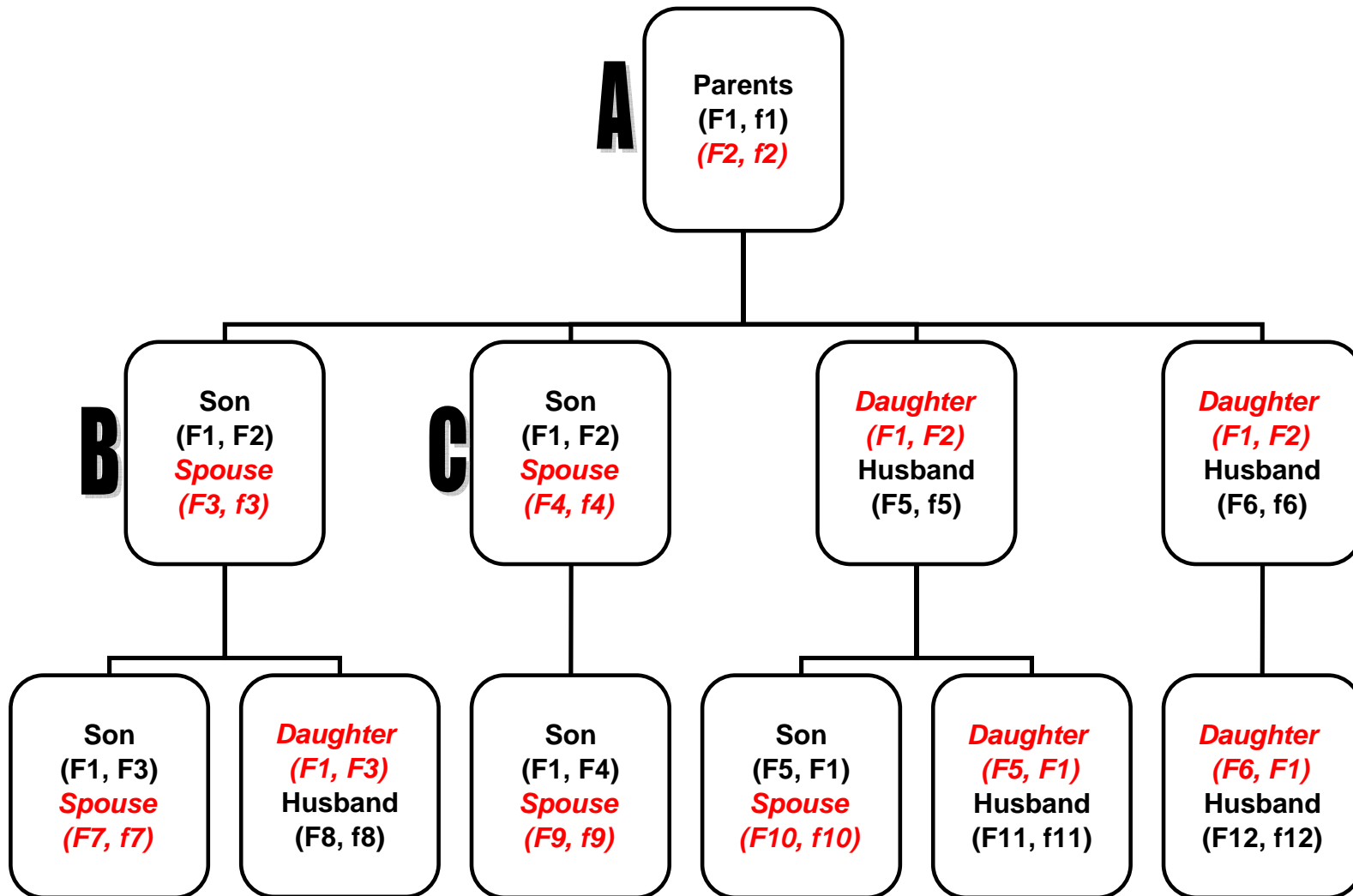
A: Village Marginality and Inequality
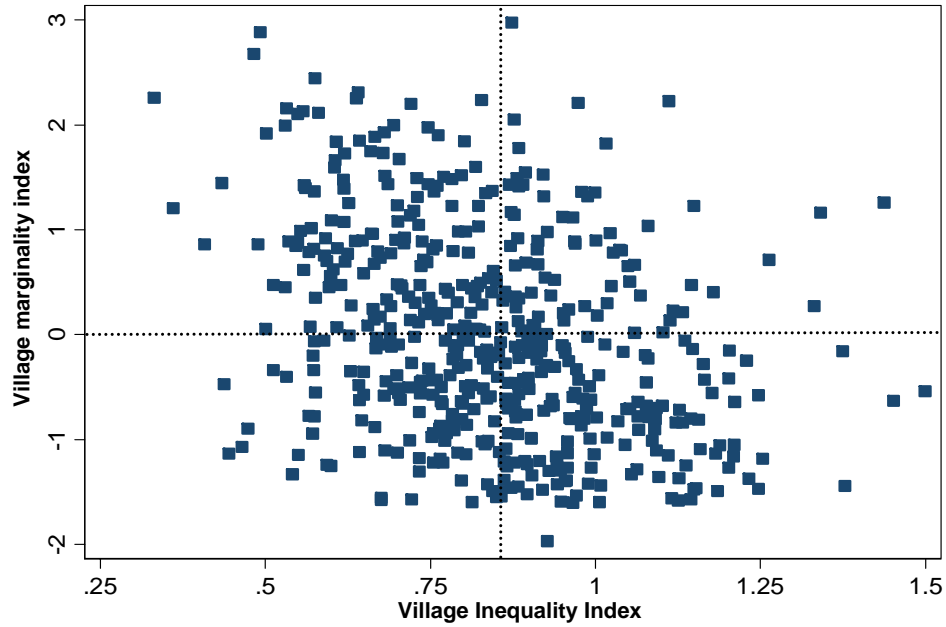
B: Village Marginality, Inequality, and Village Size

- ■ Village marginality index
- ○ Village inequality index
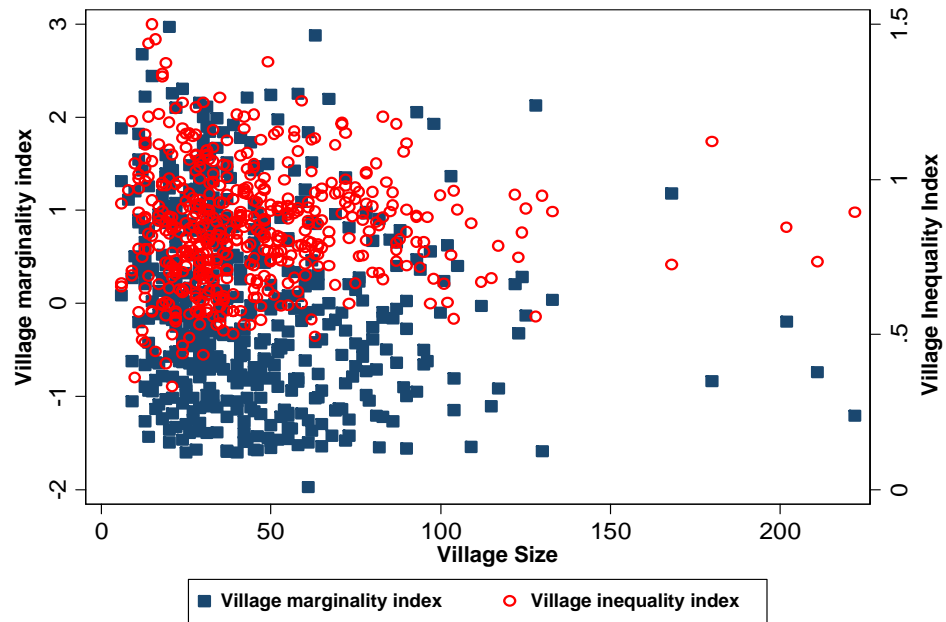
Notes: In each figure, there is one observation at the village level. The village marginality index is constructed from information on the share of illiterate adults in the village, the share of dwellings without water, drainage systems, electricity, and with floors of dirt, the average number of occupants per room in village households, the share of the population working in the primary sector, distances from other villages, and health and school infrastructures located in the village. A higher marginality index corresponds to the village being more marginal (poorer). The household welfare index is a weighted average of household income (excluding children), household size, durables, land and livestock, education, and other physical characteristics of the dwelling. The index is designed to give relatively greater weight to correlates of permanent income rather than current income. An increase in the index implies the households is less poor. The measure of village inequality is the standard deviation of the welfare index of all households in the village. Village size is defined as the number of households in the village. The village marginality index is standardized across all villages. The household welfare index is defined relative to a state norm. Hence the village inequality index is standardized at the state level. Hence the village marginality index and standard deviation of household welfare at the village level are both standardized at the regional level. Two villages are dropped (of size two and eleven) in which there are no households with any extended family links to others in the same village

Figure 2: Family Tree

A **Parents**
(F1, f1)
(F2, f2)

B **Son**
(F1, F2)
Spouse
(F3, f3)

C **Son**
(F1, F2)
Spouse
(F4, f4)

Daughter
(F1, F2)
Husband
(F5, f5)

Daughter
(F1, F2)
Husband
(F6, f6)

Son
(F1, F3)
Spouse
(F7, f7)

Daughter
(F1, F3)
Husband
(F8, f8)

Son
(F1, F4)
Spouse
(F9, f9)

Son
(F5, F1)
Spouse
(F10, f10)

Daughter
(F5, F1)
Husband
(F11, f11)

Daughter
(F6, F1)
Husband
(F12, f12)

Notes: We use the convention that the head's surnames are written in standard (black) font, and those of his wife are written in (red) italics. Paternal surnames are indicated in upper case (F1, *F2*) and maternal surnames are indicated in lower case (f1, *f2*). First names are not shown as they are not relevant for the construction of extended family ties. Each household in the family tree is assumed to be couple headed purely to ease the exposition.
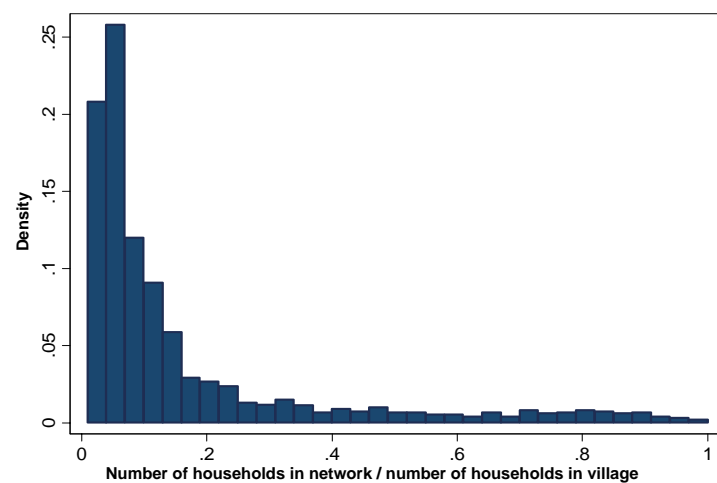
Figure 3: Family Network Descriptives

A. Number of Family Networks in the Village
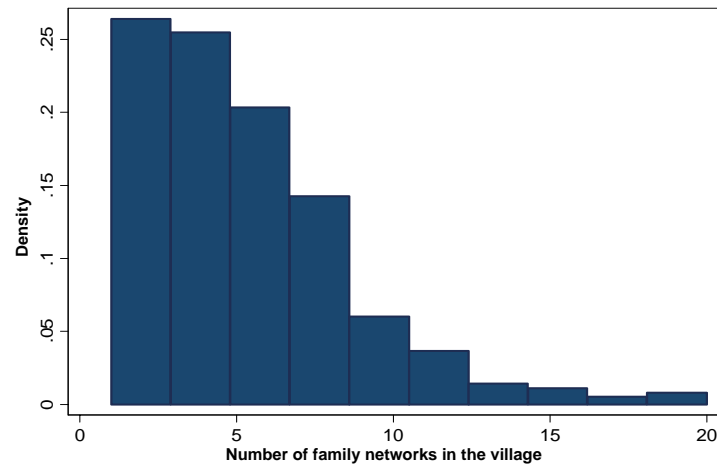
### B. Size Distribution of Family Networks
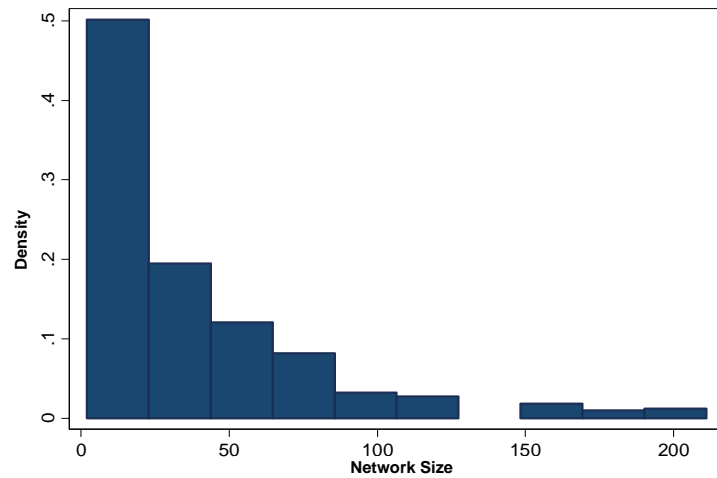
### C. Family Network Size as a Ratio of Number of Households in the Village

Notes: Each figure is constructed from those family networks with at least two households in them. Of the baseline sample of 22553 households that can be tracked over the first and third waves of Progresa, 17030 (75.5%) of them are within family networks with at least two households. There are 2196 family networks in total.

Figure 4: Family Network Graphs, at Median Village Size

A. Disperse Village
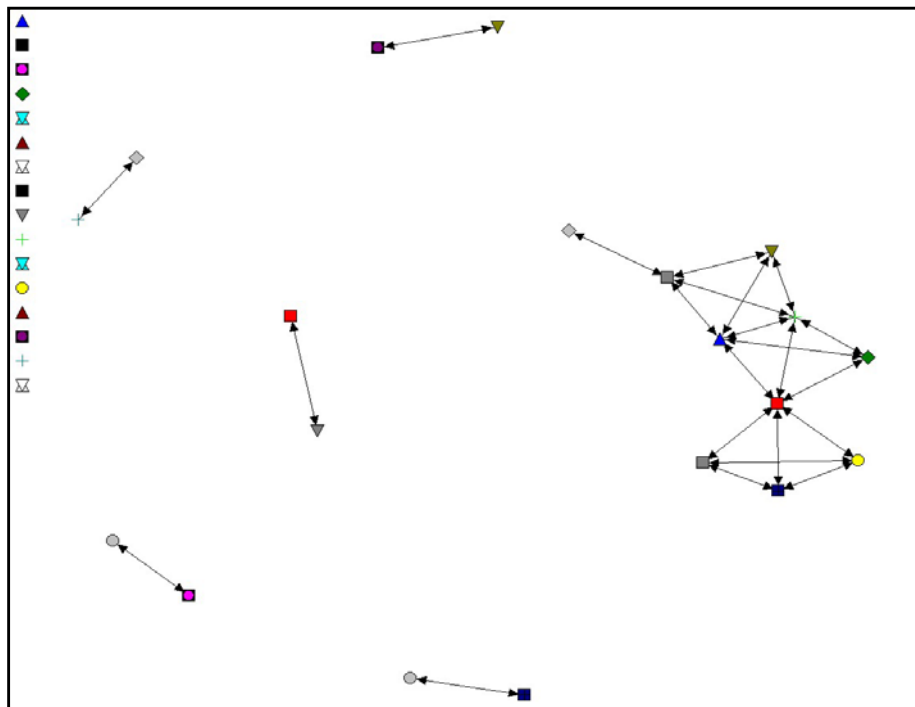
B. Interconnected Village
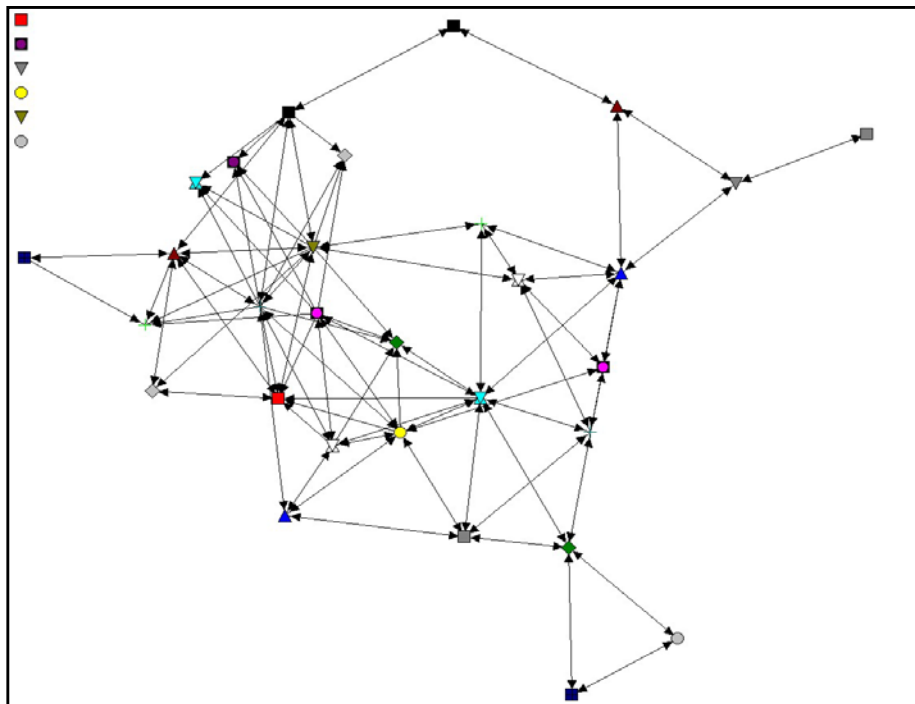
Notes: The two villages shown in Figures A and B have the same number of households in them. The number of households in each is 36, which is the median village size in the Progresa data. Each node represents a household. Each link between households correspond either to a parent/child link, a child/parent link, or a sibling link. Single node households that are not linked to any other households are shown in the top left hand corner of each graph. The figures are generated using UCINET.