INSURANCE IN THE EXTENDED FAMILY*

MANUELA ANGELUCCI† GIACOMO DE GIORGI‡

MARCOS A.RANGEL§ IMRAN RASUL¶

First Version: February, 1 2007

This version: November 11, 2008

Preliminary and incomplete - please do not quote

Abstract

We study the role of extended family networks as providers of insurance and investment in marginalized rural Mexican villages. Our main conjectures are: 1) the family achieves high insurance levels because its members have limited information, enforcement, and commitment issues; 2) insurance leads to higher investment, if it has diminishing marginal returns or lumpiness. We label households with and without relatives in the village "connected" and "isolated". Despite having similar characteristics in 1997, the connected have a smoother consumption and higher consumption, income, and investment than the isolated. Households share risk only within the extended family and a large share of connected households fully share idiosyncratic risk. When hit by negative health shocks, the isolated increase child labor, reduce school attendance, and deplete their livestock more than the connected. Positive income shocks increase education only for the connected. We conclude that the extended family provides insurance and favors human capital investment.

*The paper has been screened to ensure no confidential information is revealed. We thank Joe Altonji, Pat Bayer, Raj Chetty, Luigi Pistaferri, Mark Rosenzweig, Adam Szeidl, Chris Udry, and seminar participants at Duke, Yale, Stanford, and UCSD, among others, for their comments. All errors remain our own.

†Department of Economics, University of Arizona [angelucm@eller.arizona.edu].

‡Department of Economics, Stanford University [degiorgi@stanford.edu].

§Harris School of Public Policy, University of Chicago [rangelm@uchicago.edu].

¶Department of Economics, University College London [i.rasul@ucl.ac.uk].

1 Introduction

This paper has two goals: to analyze whether and to what extent extended family networks enable their members to smooth consumption, and to study the link between insurance and investment.

Poor households in developing countries face high income volatility but have limited access to formal insurance and credit markets, so resort to informal risk-sharing arrangements. Repeated interactions between small groups of agents may be an important mean to address the information and enforcement problems that often limit the amount of informal insurance. These facts suggest the extended family might be an important channel to insure against risk: individuals know each other well and have access to information; altruism and evolutionary biology (e.g. the need to transmit and protect one's genes) may contribute to the enforcement of insurance agreements.

Consumption smoothing is tightly linked to investment. Poor households are probably willing to engage in very costly activities to ensure their consumption never falls below subsistence levels, including cutting on high-return investments. The need to disinvest to face short-term income fluctuations might have long-term negative consequences. A households that is forced to withdraw its children from school and deplete its assets to face a negative shock is deemed to live in permanent poverty, compared to a households that can, e.g., borrow and pay the loan after its investments (in human or physical capital) have become profitable. Thus, the provision of insurance has a direct effect on investment. However, while uninsured households may need to resort to costly disinvestments to smooth consumption in the presence of negative shocks, they may invest more than insured households when they have a positive shock, as they don't have to share their resources with others, for instance.

Since insurance affects both the smoothness of consumption and the investment prospects, an important metric to understand the comparative importance of the family is not only the relative smoothness of consumption for the connected and the isolated, but also how different their investment strategies are.

Understanding whether and to what extent members of extended family networks can insure each other, and how this affects investment decisions and outcomes has important implications in terms of design of optimal policies. For example, it helps identify vulnerable households that are not well-insured against negative income shock. More broadly, it sheds light on the role of informal institutions as means to overcome market imperfections in developing countries.

Our analysis focuses on a set of marginalized villages in rural Mexico and uses the data collected for the Progresa experiment, as well as subsequent waves after the completion of the experiment. As such, we are looking at a population with a high need for insurance but no access to financial markets, the obvious target of various forms of policy interventions.

We model the economic agents as making random income draws and being able to smooth consumption through transfers. However, transfers are costly. The transaction cost, born by the agent making the transfer, is a reduced form term for all information and enforcement problems that may limit the degree of insurance. The existence of transaction costs limits the amount of insurance between groups of agents. When the cost exceeds a certain threshold, the agents will not insure each other against transitory income shocks. Since transaction costs are likely lower for related than for unrelated households, we expect related households to have larger transfers than unrelated households in any state of the world, at the very least, and only the former group to have any transfer at all, in a more extreme case.

When the agents can invest, they will do so depending on their income, the transfer from the other agent, and on whether the investment is hidden or observable. TBC

Our empirical analysis begins by describing the characteristics of households with and without an extended family network in the village, which we call "connected" and "isolated". We are interested in identifying the main correlates of connectedness, understanding whether connected and isolated households have different need and preference for insurance, and comparing their changes in income, investment, and consumption.

Our next step is establishing whether households share risk only within the extended family or also with other villagers. Our data enable us to test whether, when a given household has an exogenous income increases, consumption, transfers, and informal loans of its relatives and non-relatives increase. If households share risk only within the extended family, then consumption and net transfers or loans should increase only for related households and should be a positive function of the exogenous income shock.

After having established that households share risk only within the extended family, as we will describe below, we document the extent of risk sharing achieved within the family network and compare the variance of consumption for the connected and the isolated.

Lastly, we test the hypothesis that the provision of insurance increases investment in human capital by comparing the effect of negative and positive shocks on human capital investments for connected and isolated households.

Our analysis exploits three unique features of the data. First, we have a census of 506 marginalized villages in rural Mexico. This is a panel of about 20,000 households interviewed 8 times between 1997 and 2003, providing information on consumption, income, health and weather shocks, and informal loans and gifts, as well as other socio-demographic characteristics. Second, we can identify the extended family in each village by matching household heads' and spouses' last names, exploiting the Hispanic convention of having two surnames, the first from the paternal lineage and the second from the maternal one. Third, between 1998 and 1999 Progresa is offered only in a random group of 320 villages, in which poor households are eligible for sizeable cash transfers, while no public transfers takes place in the control villages.

Our main results can be summarized as follows. Beginning with the characteristics of connected and isolated households, we establish that at least 80% of the households in our villages are part of an extended family network. Connected and isolated households are remarkably similar on average characteristics, in particular when we drop single-headed households. They have identical consumption, income, wealth index, and employment, and similar asset ownership. The main differences between the connected and isolated are attributable to life cycle effects. Isolated household heads are 2.5 years older, hence they have fewer young children, a lower literacy rate (which is higher for more recent cohorts), more irrigated land and poultry, but lower non-irrigated land holdings and lower rates of seasonal migration to the US. This preliminary analysis suggest, among other things, that connected and isolated households may face similar amounts of risk and need for insurance.

While similar at baseline, income, consumption, and investment change differently for connected and isolated households over time. A difference-in-difference comparison of these variables shows that the change in income, consumption, and investment and significantly higher for the connected. This preliminary evidence is consistent with our hypothesis that the extended family, by providing insurance, enables its members to undertake more investment, increasing their income and consumption.

Proceeding to study the relationship between family networks and insurance, we establish one important result: connected households share risk within their family network but not with isolated households. When eligible households receive the Progresa grants, their ineligible relatives consume more and receive more informal loans. At the same time consumption and transfers to ineligibles who have no eligible relative (or no relative at all) in the village do not. We also show that insurance seems related the "connectedness" of the households, rather than other features. That it, insurance does not seem related to ethnicity or land ownership, and these variables are not associated with higher degrees of insurance over and above the one associated with connectedness. The relevant institution, therefore, is the extended family, rather than one's ethnicity or occupational group.

Consistent with the model predictions, consumption for the ineligibles is a positive function of the size of the income shock per network members, i.e. the more money is injected in the network, the more its members will benefit, irrespective of which household receives the extra income. On average, for each extra dollar received by an eligible individual, her non-durable consumption increases by 54 cents and the consumption of her ineligible relatives increases by 13 cents. The consumption of eligibles who have no ineligible relatives increases by 69 cents for each dollar received.

There is a high degree of insurance among the extended family. The Mace/Townsend regressions suggest that connected household achieve full risk sharing in 80% of the networks (at the 90% confidence level), 10 to 15% more often than for the isolated. [TBC: check this is true]

As the isolated do not share risk within the village, they may be more vulnerable to negative idiosyncratic shocks. Comparing the longitudinal variance of consumption and income shows that both are smoother for connected than isolated households (by 6 and 2% respectively, if we include single-headed households). However, this in itself is not necessarily evidence that the isolated are less insured than the connected, or that they engage is costlier activities to smooth consumption. Moreover, consumption is much smoother than income for both groups, as the ratio of the longitudinal coefficient of variation of income and consumption is XX for the connected and YY for the isolated. Therefore, we compare the effect of idiosyncratic shocks on child labor supply and school enrollment, livestock, and assets. In this way we can test whether the cost of smoothing consumption is higher for the isolated than for the connected, when hit by a negative shock, and which group invests more in the presence of good shocks. We find that isolated household respond to negative shocks with larger decreases in school enrollment, increases in child labor, and depletion of livestock (poultry and oxen) than connected households. Therefore, consumption smoothing is costlier for the isolated, who, by reducing investment, are foregoing future wealth in order to stabilize current consumption.

The paper has the following structure: in Section 2 we sketch a simple risk-sharing model to define the notation, show the standard full insurance results, and prove that full insurance cannot be achieved with positive transaction costs, which are probably higher between isolated than connected households. Sections 3 and 4 describe the data, the creation of family networks and their characteristics. Section 5 shows that risk-sharing occurs only within the extended family; Section 6 shows that most families are fully insured against idiosyncratic risk, while Section 7 documents that isolated households engage in costlier activities to smooth consumption. Section 8 concludes.

2 A model of partial insurance and investment

2.1 Partial insurance

Consider a simple, two-agent, two-period insurance model, with an iceberg-type transaction cost such that part of the transferred resources melt away in the transfer. This cost captures all the potential reasons - e.g. imperfect information, limited commitment - why full insurance may not be achieved.

We conjecture that members of an extended family face lower, possibly zero, costs of transacting among each others than with households outside their family ties. This is because family members are part of a small network (e.g. Genicot and Ray 2003) with repeated interactions and potentially high trust (Mobius and Szeidl 2007) and altruism (Altonji et al., 1992; Foster and Rosenzweig, 2001), features associated with stable networks achieving high levels of insurance. That is, the extended family may provide low-cost information about its members' income and effort and reduce the commitment problems, achieving a higher degree of insurance than unrelated households. Some of the latter households may have no insurance at all, if the transaction cost is sufficiently high.

Consider two risk averse agents, $i \in \{1, 2\}$, with instantaneous utility function $u_i = (1 - \delta)\ln(c_i)$, a rate of inter-temporal preference $\delta < 1$, an endowment y_i, no storage technology, and no leisure. Each agent receives a random income $y_i(s) = \bar{y}_i + \epsilon_i(s)$ that varies in different (finite) states of the world $s \in S$ and all risk is idiosyncratic. The income is the sum of a deterministic component, $\bar{y}_i$, and a random shock $\epsilon_i(s)$. We assume that the distributions of pairs of random shocks $(\epsilon_1(s), \epsilon_2(s))$ is symmetric and independent over time, as in Coate and Ravallion (1993). Each state of the world occurs with a positive probability $\pi(s)$, with $\sum_s \pi(s) = 1$. The agents smooth consumption making transfers to each other.

In the absence of transaction costs, the sum of the agents' consumption equals the total available resources in each state of the world, $c_1(s) + c_2(s) = y_1(s) + y_2(s) > 0$. With positive transaction costs a fixed proportion of the resources transferred is destroyed. Thus, a transfer d to agent 1 requires $(1+\alpha)d$ transfer from agent 2 when $\alpha > 0$. In this case, total consumption is lower than aggregate resources, $c_1(s) + c_2(s) = y_1(s) + y_2(s) - \alpha d(s)$ if $(s) \neq 0$.

The social planner maximizes a weighted average of the agents' utility functions, where the weights $\lambda \in (0;1)$ and $(1-\lambda)$, suppose for now that $\lambda = .5$ represent the relative importance of the two agents. Consider the case in which agent 2 makes a transfer to agent 1, implicitly we are assuming that $y_2 > y_1$. The maximization problem for the social planner is (time subscript omitted):

$$
\begin{aligned}
& max_{c_1,c_2,d} U = \sum_s \pi(s) \left[\ln c_1(s) + \ln c_2(s)\right] \\
& s.t. \\
& c_1(s) = y_1(s) + d(s) \\
& c_2(s) = y_2(s) - (1+\alpha)d(s) \\
& y_1(s) + y_2(s) = C(s) + \alpha d(s) \\
& c_1(s), c_2(s) > 0; d(s) \geq 0
\end{aligned}
$$

From the first-order conditions we can derive the optimal consumption levels in each state.

These are:

$$c_1^*(s) = \frac{1}{2(1+\alpha)}[Y(s) + \alpha y_1] \tag{1}$$

$$c_2^*(s) = \frac{1}{2}[(1+\alpha)Y(s) - \alpha y_2] \tag{2}$$

$$d^*(s) = \frac{y_2 - (1+\alpha)y_1}{2(1+\alpha)} \tag{3}$$

$$\tag{4}$$

It is easy to show that the amount transferred depends negatively on α and that households consumption is now a positive function of household income. Therefore in the presence of positive transaction costs the longitudinal variance of consumption is higher for each household, given that perfect insurance cannot be achieved. In the absence of transaction costs, i.e. when $\alpha = 0$, the agents can fully insure against idiosyncratic risk. Therefore consumption and its variance depend only on aggregate resources and the Pareto weights, and not on the individual endowments. An increase in agent 2's income, Δy_2, increases her consumption by $\Delta c_2^* = \frac{1}{2}\Delta y_2$ and agent 1's consumption by $\Delta c_1^* = \frac{1}{2}\Delta y_2$. In this particular example where we assumed equal Pareto weights.

When $\alpha > 0$ the effect of the transaction cost is to decrease the level (of c_1) and increase the overtime (states) variation of consumption. Individual consumption now depends on each agent's own income (through d), as well as on aggregate resources, and full insurance cannot be achieved. The higher the transaction costs (i.e. the higher α), the further away from full insurance. Depending on the state of the world, there will be a threshold for α so that $d^*(s) = 0$. Both agent's consumption increases in response to a higher income Δy_2, but this increase in consumption is smaller than with full insurance and grows inversely with transaction costs, i.e. $\Delta c_1^* = \frac{1}{2(1+\alpha)}\Delta y_2$. Part of the resources are destroyed because of the transaction costs (e.g. agents face monitoring or enforcement costs).

If α is high, depending on the income realizations, there is autarky, i.e $d^*(s) = 0$ if $\alpha = \frac{y_2(s) - y_1(s)}{y_1(s)} \equiv \bar{\alpha}$. Each agent consumes one's own endowment, makes or receives no transfers, and the longitudinal variance of one's consumption equals one's income variance. An increase in agent 2's income Δy_2 results in an increase in her consumption for the same amount, $\Delta c_2^* = \Delta y_2$ and in no change in agent 1's consumption.

We exploit these features of our model to test whether connected and isolated households 1) form separate insurance groups because they face different transaction costs and 2) achieve different levels of insurance. To answer these questions we perform a variety of tests.

<u>Hypothesis 1: households are more likely to share risk if they are connected.</u> Consider two pairs of agents 1 and 2. The first pair of agents are connected (K) and face low transaction costs $\alpha \to 0$; the second are not (I) and face high transaction costs $\alpha \to \bar{\alpha}$. Agent 2's income increases exogenously by an amount Δy_2 in both cases. For the connected pair, consumption increases for both agents. The transfer to agent 1 increases by the same amount as his consumption does. For the isolated pair, consumption increase only for agent 2, $\Delta c_2^{I*} = \Delta y_2$, while agent 1's consumption does not change and he continues to receive no transfer. Our data enable us to observe how the consumption of and transfers received by a set of "agent 1" households changes in response to an exogenous income increase for "agent 2" households depending to whether they are related to this latter group or not. Therefore, we can test the following hypotheses:

1.1 For a given income increase Δy_2, consumption and transfers will increase only for households connected to agent 2, i.e. $\Delta c_1^{K*} > 0$ and $\Delta d^{K*} > 0$, but $\Delta c_1^{I*} = 0$ and $\Delta d_1^{I*} = 0$ if the cost of transacting is high enough.

1.2 For a given income increase Δy_2, agent 2's consumption will increase more if she is isolated rather than connected, $\Delta c_2^{K*} < \Delta c_2^{I*}$.

1.3 The consumption of and transfers to connected agent 1 households increase only if they are connected to some agent 2 household whose income is increasing.

The tests above exploit the partial population experiment we observe in our data.

While for the next battery of tests we do not explicitly exploit the experimental nature of PROGRESA.

<u>Hypothesis 2: connected households have more insurance than isolated households.</u>

1.1 If $\alpha = 0$, the allocation of consumption will be the first best $c_1^*(s) = c_2^*(s) = \frac{Y(s)}{2}$. While with $\alpha > 0$ the allocation will move further away from first best, with increasing α up to the autarkic allocation for $\alpha = \bar{\alpha}$.

For our next set of tests, consider the first-order conditions in the absence of transaction costs, i.e. when $\alpha = 0$. Taking the log of equation (1) and then the first difference yields a well-known relationship between the growth of individual and aggregate consumption for household i at time t:

$$\Delta lnc^*_{ht} = \Delta ln\bar{C}_t$$

Therefore, idyosincratic variables should not enter such regression. We can then regress the growth in household consumption over the growth in aggregate consumption and household income,

$$\Delta lnc_{ht} = \beta_1 \Delta lnC_t + \beta_2 \Delta lny_{ht} + u_{ht} \tag{5}$$

where Δlny_{ht} is the growth in income for household h at time t and $u_{ht} = \Delta\epsilon_{ht}$ is an error term derived from assuming a multiplicative error in the measurement of consumption, $c_{ht} = c^*_{ht}e^{\epsilon_{ht}}$. We use equation (5) for different purposes. First, we estimate this regression for all connected and isolated households pooled, and test jointly whether $\beta_1 = 1$ and $\beta_2 = 0$. Although, β_1 mechanically goes to 1 making the important component of the test that based on β_2. Since we reject the null of full-insurance in the aggregate, we proceed by testing it separately for connected and isolated households in the entire sample and then at the village level.

2.2 Insurance and investment

Before performing these tests we have to briefly review some features of the program and of the evaluation design, explain how we create family networks, and describe the networks' and villages' main features.

3 Progresa and its evaluation

Progresa is an anti-poverty program that targets poor households in rural Mexico. The average transfer is 200 pesos, equivalent to 22% of eligible households' monthly income (De Jainvry and Sadoulet, 2006) and to 25% of pre-program household's food consumption, which averages 160 pesos for adult (equivalent) in recipients households (Angelucci and De Giorgi 2007). About 75% of households are classified as eligible based on their poverty status as computed October 1997, although only a smaller share are treated initially.[1]

[1]The initial allocations of households between eligible and ineligible that placed 50% of the households in the eligible group was subsequently revised just before the roll out of the program; however, in the first year of the

While the main grant component is conditional on children attending classes between 3rd and 9th grade and having periodic health checks, the conditionality is actually binding only for a subset of households. This is because pre-program enrollment rates, up to 6th grade, are higher than 90% and because the smaller health and nutrition monetary transfer is only conditional on infrequent health checks. Only the transfer for 7th to 9th grade attendance is actually conditional for a larger group of families, as pre-program attendance for these grades is only about 60%. Thus, the Progresa transfer has a pure income effect for most recipients.

At the time Progresa started, now Oportunidades, there were several other aid programs in the area. Progresa was however engineered to replace all the pre-existing (highly ineffective) programs.

Our data have three important characteristics. First, they provide a complete census of all households from 506 rural villages in 7 different states.[2] The data are longitudinal, with an initial wave collected in October 1997 and then approximately every 6 months until November 2000, and lastly in November 2003. We have income data in all cross sections, and consumption data in 6 cross sections: March 1998, right before the program start, November 1998, May and November 1999, November 2000 and 2003. We have complete data for about 22,500 households per wave up to November 1998, about 20,000 up to November 2000, and 19,000 in 2003.

Second, between May 1998 and November 1999 the program is offered only to a random group of 320 villages. The remaining 186 villages receive their first transfers only at the end of 1999. Therefore we have three data waves, November 1998, May 1999, and November 1999, in which only a random subset of our sample of villages was treated.[3]

Since the data are a census, we have information on both eligible and non-eligible households, and we can group households according to their poverty status (which defines eligibility) and village of residence, as shown in Figure **??**. This figure shows our data provide a partial-population experiment (Moffitt, 2001), as the treatment is offered only to a subset of the villagers. Thus, we can measure the effect of the Progresa transfers to eligible households on both consumption and transfers to ineligible households.

Third, we are able to construct family links and therefore extended families. In the next

program most of the re-classified households, typically elderly, did not receive any transfer for administrative reasons. In our analysis we keep those observations as treated.

[2]Guerrero, Hidalgo, Michocoan, Puebla, Queretaro, San-Luis Potosi and Veracruz; mostly in central Mexico.

[3]In fact in November 1999 some small fraction of control eligibles started to receive some transfers. If those transfers are not negligible we might be estimating a lower bound to true ITE.

section we describe how we identify these related households and how we generate family networks.

4 Family networks in rural Mexico

4.1 Network creation

To construct extended family links between households in the same village we exploit information on surnames provided in the third wave of data and the fact that 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). Hence a household that is headed by a husband and wife, has four associated surnames – the paternal and maternal surname of the head, and the paternal and maternal surname of his wife.

We use these data to create within-village family networks, starting from degree one, i.e. each household head and spouse's parents, offspring, and siblings. We then proceed further to get at the extended family networks, i.e. grand-parents, aunts and uncles, etc.. All family links are defined across households on the basis of *two* surname matches. We believe our matching algorithm to be reliable. For example, by definition a household cannot have parental links to more than two other households (the parent's of the head and the parent's of the spouse), conditional on the parents not being present within the household. Reassuringly, this is true for 97% of households. Moreover, a comparison with network data from the *Mexican Family Life Survey* (MxFLS), collected in 2001, shows that the within-village links we identify in our data are indeed no larger than those measured in the MxFLS for a comparable sub-sample of households using relatives from *any* location. Angelucci, De Giorgi, Rangel, and Rasul (2006, 2007) discuss the creation of these family networks, the characteristics of the matching algorithm, and the possibility and extent of measurement error in great details.

Table 1 provides a brief description of our family networks. There are 7.8 households within the same family network. The average distance between any two households in the same network is 2.12. The degree of the average network, defined as the number of households each is directly connected to, is 2.07. The diameter of the network, defined as the largest distance between any two households in the network, is around 2.5 on average. Hence family networks are unlikely to span across more than three generations.

Approximately 40% of the networks are composed by 2 households. About 50% of all networks are a mix of eligible and ineligible households (and about two thirds of networks with at least three households).

4.2 Correlates of Extended Family Links

Families form and dissolve through the marriage, death, or migration of their members. These processes are unlikely random and some of them might be directly related to insurance motives (e.g. Rosenzweig and Stark 1989). It is therefore useful to establish how comparable connected and isolated households are in terms of observable characteristics. This comparison also helps us understand the possible econometric concerns that will arise for the later analysis.

TBC: add means of weather and health shocks

Table 2 shows the means of key demographic and socioeconomic variables from the 1997 data. Many of the apparent differences between connected and isolated households emerging from the table right panel depend in fact on the different shares of single-headed households, who are poorer, less educated, and larger than couple-headed households. A fraction of single-headed households are likely erroneously classified as isolated because we do not observe the links of the missing spouse (although those links may be inactive if the direct relative is absent or dead) or adult children who live in a separate household. Once we drop single-headed households and compare both demographic and economic outcomes, which we do in the left panel, connected and isolated households appear remarkably similar.

The main differences between the two groups is that isolated households heads are 2.5 years older, have fewer young children, and a 9% higher share of illiteracy than connected heads. Most of the other variables do not differ between the connected and isolated. In particular, income, food and non-food expenditure, and wealth index are almost identical, as well as the total amount of land held, adult and child labor, and most assets and livestock. The few exceptions are that the isolated are more likely to own irrigated land and hold a larger stock of chickens, but have fewer temporary (but not permanent) US migrants and own fewer stoves and TV sets.

We also performed a Pearson's chi squared test of the difference in the distribution of unemployment and employment type for all members of connected and isolated households at least 8 years old. Different employment patterns may reflect different risk preferences or need for insurance. The main labor force categories are unemployed (68%), daily worker (15%), non-agricultural manual worker (4.5%), self-employed without (4.3%) and with personnel (0.12%), working for a family business without pay (4.4%), and ejidatario (2.5%). We find that no significant difference in the employment type for isolated and connected households.

The correlation between the age of the head and spouse of a household and the likelihood that their parents, adult children and siblings live in the same village in a separate household is somewhat mechanical. Angelucci et al. (2007) show that the rate at which a household head "loses" links as he ages (i.e. parents passing away) is faster than the rate at which he "acquires" new links (through children being born). This makes households with older heads more likely to be isolated.

Thus, these difference seem overall attributable to life cycle and cohort effects: isolated households are older, therefore its adult members are less likely to be international migrants, and more likely to invest in livestock than in migrations; being older, they had more time to purchase "good" land. Because they are more "traditional", they own fewer TV sets and stoves.

Comparing means for connected and isolated households is somewhat misleading if there is assortative matching. For example, all the wealthiest and poorest households may belong to distinct family network, but by aggregating all families we would not detect it. We test for assortative matching by computing the standard deviation of the wealth index for each family network and each village. If there is positive (negative) assortative matching then the ratio of network over village standard deviation, $\frac{sd^n}{sd^v}$, would be less (more) than one. Figure 1 shows that these ratios are centered around one, rejecting the hypothesis of assortative matching. This confirms the previous conclusion, namely that one cannot easily predict whether a household is connected or isolated by looking at its observable characteristics.

TBC: add correlation for health shocks

Table 3 shows the correlation in the occurrence of natural disasters for two members of the same extended family dynasty (column 1) and for two random households in the village (column 2). These correlations are not systematically different for the two types of households, suggesting the following: first, the desire for insurance against weather shocks does not seem to be the driving force behind family formation, in which case we would expect the correlations to be lower for family members; second, this confirms the earlier evidence that isolated and connected households have similar observable characteristics.

4.3 Variances and changes in consumption and income

To complete this preliminary analysis of the characteristics of connected and isolated households, we compare the variance of consumption and income and their differences between 1997 and 2003.

Table 4 shows that consumption is more dispersed for isolated households. Consider food consumption, which is measured more accurately. Both its cross sectional and the longitudinal variation are larger for the isolated, by 25% and 10% for all households. Figure 2, which plots the time profiles of food and income for a random connected and isolated households in randomly drawn villages with similar characteristics, confirms these findings.

If we take a measure of aggregate food consumption variance as the variation of village or network consumption over time, i.e. aggregate uncertainty, we find a similar pattern: connected households appear to smooth idiosyncratic shocks better than isolated ones. However, the longitudinal variance of income is roughly 4% smaller for connected households. In the econometric analysis we will test whether consumption is smoother for the connected conditional on their income variance.

****

Questions about this table

1) should we drop the part with non-food consumption from the table, since the CVs are the same for K and I?

2) so far we have tested for significance diff. between K and I. Doesn't it make sense to also test for significant differences in the CV here? Or is it ok to leave it to later? It looks like a waste of time (and a logical inconsistency) not to do the test here.

3) what's the point of doing hh CV/village cv? the denominator is the same for K and I so what does it add compared to the previous row?

4) if we leave both food and non-food consumption we have to explain why income is so much lower than consumption (food + nonfood)

5) which variable did we use for total income? In any case, we have to figure out why it is so small. (esp if it includes earnings from sales of goods, grains, etc). This could actually be a

good motivation for why we DO NOT want to rely on income too much.

****

Table 5 compares the difference in income, consumption, and investment levels for connected and isolated households between 1997 and 2003. Provided that the connected households are more insured than the isolated, as we show below, if there is a positive link between insurance and investment we expect to find higher investment for the connected than for the isolated in 2003, conditional on their 1997 levels. If investment causes long-lasting income increases, this will also be reflected in higher consumption.

The descriptive evidence is consistent with this conjecture: the change in investment - proxied by the share of household members with at least secondary education and by the stock of animals (which are both store of value and productive assets) and machinery is significantly lower for the isolated than for the connected. Income and consumption (especially durable assets) are also significantly lower for the isolated.

5 Do family networks form insurance sub-groups?

5.1 Families and insurance subgroups: exploiting the program randomization

This set of tests exploits the random variation in income resulting from the experimental design of Progresa. For the first 18 months of its existence Progresa was implemented only in a subset of the sampled villages, the policy can be seen as an exogenous income shock to "agent 2" households (the eligibles), in our model, on the transfers/loans to and consumption of different groups of households. If we abstract from its effect on secondary school enrollment, which increased by about 10 percentage points among eligible households without affecting the enrollment rates in ineligible households (Angelucci *et al.*, 2007), the experimental variation we observe in the data fits our model. In particular, we observe that a large part of the Progresa transfer comes as a windfall transfer, given that pre-program enrolment rates were already above 90% for primary school and above 65% for secondary school. As such, income increases exogenously for eligible households and we can observe how consumption changes for different groups of households depending on their relationship to the eligibles.

Although it has some unique features that distinguish it from the previous poverty alleviation programs, Progresa is just the last of of several aid programs targeting the poor in rural Mexico. The recipients are used to receiving government assistance of different forms; the possibility of such policies should then be considered part of the information set of those agents. One can think of Progresa as one of the states of the world that households write implicit contracts upon, rather than an entirely new occurrence. Further, since there was uncertainty as to whether the program would continue beyond 1999, households in treatment villages likely considered it as a temporary income shock that would not change their relative power in the (implicit) contract, i.e. the Pareto weights.

Since our tests estimate treatment effects on ineligible and eligible households, we briefly define these parameters and discuss their identification. Define Y_{1i} and Y_{0i} as the potential outcomes for household i in treatment villages ($P_i = 1$) *in the presence* and *in the absence* of the treatment, where the treatment is the existence of Progresa transfers to eligible households ($N_i = 0$) in treatment villages ($P_i = 1$). We call the average effects of the program on 1) eligible households and 2) ineligible households ($N_i = 1$) living in treatment villages, Average Treatment Effect on the Eligibles (ATE) and Indirect Treatment Effect (ITE), and we define them as follows:

$$\begin{aligned} ATE(Y)^f &= E(Y_{1i}|P_i = 1, N_i = 0, f = 1) - E(Y_{0i}|P_i = 1, N_i = 0, f = 1) \\ ITE(Y)^f &= E(Y_{1i}|P_i = 1, N_i = 1, f = 1) - E(Y_{0i}|P_i = 1, N_i = 1, f = 1) \\ & \qquad f = \{K, I\} \end{aligned}$$

Under the assumptions of random assignment and in the absence of program spillover effects to control villages, the expected value of the potential outcome in the absence of the treatment, Y_0, is the same in both treatment and control villages, i.e. $E(Y_{0i}|P_i = 1, N_i = j, f = 1) = E(Y_{0i}|P_i = 0, N_i = j, f = 1)$, for $j = \{0, 1\}$ and $f = \{K, I\}$. Therefore, the differences

$$ATE(Y)^f = E(Y_i|P_i = 1, N_i = 0, f = 1) - E(Y_i|P_i = 0, N_i = 0, f = 1) \tag{6}$$

$$ITE(Y)^f = E(Y_i|P_i = 1, N_i = 1, f = 1) - E(Y_i|P_i = 0, N_i = 1, f = 1) \tag{7}$$

identify the ATE and the ITE. The eligibles and the ineligibles in treatment villages are the "agent 2" and "agent 1" households from the model, and the Δy_2 is the Progresa transfer.

Therefore, we can re-express our first set of testable hypotheses using the treatment effects jargon.

1.1 If the connected households (K) share risk only between family members and not with the isolated (I), only connected ineligible households will increase consumption and transfers received because of Progresa, i.e. $ITE(c)^K > 0$ and $ITE(d)^K > 0$, but $ITE(c)^I = 0$ and $ITE(d)^I = 0$.

1.2 Among connected ineligible households, the ITEs on consumption and transfers received are positive only if these households are connected to some eligible household.

1.3 For a given Progresa transfer and similar Pareto weights, the treatment effect on consumption for connected eligible households is lower if these households are related to the ineligible with whom they share part of their transfer, i.e. $ATE(c)^K < ATE(c)^I$.

Our measures of consumption are monthly food consumption and all non-durable consumption per adult equivalent. The first variable is measured very accurately, as we have data on consumption of 37 different food items in the previous week and we observe food purchased, donated, and consumed from own production. Non-food consumption is measured with less accuracy: the recall period is longer, either one or six months, and there is a sizeable proportion of households with zero expenditure in several non-food items, as the purchase of those commodities is infrequent for indigent families. Moreover, food consumption accounts for a very large share of total consumption (about 72% for the whole sample). For these reasons, and since Progresa does not change non-food consumption of ineligible households (Angelucci and De Giorgi 2007), we focus on food consumption when estimating treatment effects for the ineligibles.

We have two measures of transfers: monetary transfers and loans received in the previous six months, and of monetary and in kind transfers from family and friends during the previous month. Both transfers and credit are informal: all transfers and 70% of loans occur among friends or relatives. Unfortunately we cannot distinguish between relatives and friends in either case, nor observe the identity of lenders and donors or whether they belong to poor or non-poor households. Second, while in principle we also have data on transfers given, this variable is unreliable, hence we cannot build a good net transfers variable. For example, in the November

1999 data 319 households report they *received* transfers from families living in the same village, while only 41 households appear to have *made* a transfer to a family in the same village. We suspect the poor may be afraid to admit they are sharing the Progresa grants with the non-poor.[4] Lastly, we observe both loans and transfers only in November 1998. In the remaining waves, we observe loans in May 1999, and transfers in November 1999.

5.2 The effect on consumption for ineligible households

In this section we test our first hypothesis that connected households do not share risk with isolated households.

We now proceed to test whether the ITEs for consumption and credit resources (transfers and loans) are larger for connected than for isolated households. We do so by estimating the following regressions:

The first empirical issue we face is whether to estimate the ITE on consumption by difference-in-difference or using only data from after the program start. While we do not observe pre-program consumption, we have a noisy measure of expenditures in March 1998. There are both conceptual and practical concerns against using this variable. At the conceptual level, while the first transfers were received after April 1998, household behavior is already affected by the treatment by March 1998. For example, in order to receive the Progresa scholarships, eligible children must go to school. That is, the payment is conditional upon verification that the eligible households complied with the program rules in the two previous months. While we are looking at the behavior of ineligible households, which do not directly benefit from the program, it is possible that these households may have some indirect benefits from the program. For example, the availability of textbooks purchased in eligible households my reduce the cost of school attendance for related households, if the related pupils share the books.[5] The second set of concerns related to the use of the March 1998 expenditure data is that 1) consumption and expenditure differ largely for these households. For example, XX, ZZ, and YY% of ineligible households report consuming home-produced corn, tortillas, and chicken, the main staples of their diet; 2) households are asked to report their expenditures for broad food

[4] We use this variable to compute net transfers. However, we also consider gross transfers only. The results do not change.

[5] We fail to find this effect for the average ineligible households with secondary school children (Angelucci *et al.*, 2007). However, Bobonis and Finan (2008) find evidence consistent with this conjecture for a sub-sample of ineligible families.

categories (vegetables, grains, meat, processed food), so they are likely to under-report their actual expenditure. At the same time, however, it is possible that there may be some true differences in pre-program consumption for some groups of households.

We end up estimating consumption ITEs exploiting alternatively only the cross sectional, and both the cross sectional and the longitudinal variation in The data, as shown in the following equations.

$$c_{it} = \gamma_0 + \gamma_1 P_i * K_i + \gamma_2 P_i * I_i + \gamma_3 X_i + \gamma_{4t} \sum_t W_t + u_{it}$$

$$c_{it} = \alpha_0 + \alpha_1 P_i * K_i + \alpha_2 P_i * I_i + \alpha_3 T_t + \alpha_4 P_i * K_i * T_t + \alpha_5 P_i * I_i * T_t + \alpha_6 X_i + \alpha_{7t} \sum_t W_t + u_{it}$$

$$d_{it} = \beta_0 + \beta_1 P_i * K_i + \beta_2 P_i * I_i + \beta_3 X_i + \beta_{4t} \sum_t W_t + u_{it}$$

where c stands for consumption and d for credit market variables, as we will list below. The variable P is a dummy for residents of Progresa ($P = 1$) and control ($P = 0$) villages, T is a time dummy for baseline ($T = 0$) and later data waves ($T = 1$), and K and I are dummies for connected ($K = 1$) and isolated ($I = 1$) households. X is a set of household and village characteristics measured at baseline, including region fixed effects.[6] The dummies W are time dummies for each of the available data waves.

The parameters γ_1 and γ_2, and α_4 and α_5 identify the ITEs for connected and isolated households. Therefore, a test for the significance of their difference is a test that these two treatment effects are statistically different. We do not observe pre-program data for the loans and transfers variables, therefore in this case we estimate the treatment effects exploiting the cross sectional variation only. The parameters β_1 and β_2 identify the ITEs for loans and transfers.

The left panel of Table 6 shows that March 1998 expenditure for ineligible households is lower, though imprecisely estimated, in treatment villages for connected and couple-headed isolated households. However, the main results from the right panel does vary depending on how we estimate the ITEs: we find a positive and significant increase in consumption for the

[6] List control variables.

connected, but not for the isolated. The estimates ITEs for the isolated vary considerably in size and sign across different types of regressions and according to how we trim the data. Part of this imprecision depend on the smaller sample size, as only about 2000 ineligibles are isolated. On the contrary, while the size and the precision of the estimated ITE^K varies somewhat across the different specifications, they all draw a consistent picture: among the ineligibles, only connected households increase their food consumption.

If households pool resources only with members of their extended family, then we would expect the ITE on consumption to be positive only for ineligible households with eligible relatives. Therefore, we estimate the consumption ITEs for connected households with and without eligible relatives separately, as reported in Table 7. We divide households in three groups, depending on the share of their extended family members entitled to the program. We experimented with different groupings and the results were always broadly similar; therefore, we report the ones grouping households with 0, up to 30%, and more than 30% eligible relatives. We find that the ITEs on food consumption are a positive function of the share of connected households, when using only the cross sectional variation in the data. The pattern is more mixed for the difference-in-difference estimate of the ITEs. However, the treatment effects are positive and significant only for the group with a sizeable share of eligible relatives (we had similar results using lower cutoffs, e.g. 25%). Thus, overall these results confirm our conjecture, namely that households benefits indirectly from the program transfers only if they are related to program beneficiaries. It would be tempting to rationalize the small ITE for households with a small fraction of eligible relatives as evidence that households share resources only if the positive shock at the network level exceeds a certain threshold. However, since there are only about 200 ineligible households with few eligible relatives, it is not clear how much one can infer from these results.[7]

The bottom right panel of Table 7 shows that the magnitude of the ITEs is a function of network size: the ITE is positive and significant only for households with large networks (i.e. related to at least 6 families).

We continue the investigation of whether extended family members share resources by considering only treated villages. We regress household log-food consumption per adult equivalent on network log-transfer per adult equivalent. Since the actual transfer is endogenous, as its

[7] The group of connected ineligibles with no eligible relatives is also small, amounting to about 230 households.

amount depends on whether eligible households comply with the program schooling and health requirements, we use potential transfer as an instrument (computed based on the maximum completed school grade in the 1996-1997 academic year and gender of eligible children). We control for predetermined household, network, and geographic variables, including number of eligible children in the network grouped by gender and primary versus secondary school level.

The IV estimates from Table 8 show that the average transfer elasticity of food consumption for the ineligibles is 0.13, that is, an extra dollar transferred to the eligible relatives increases ineligible consumption by 13 cents in the average network, and by 20 cents in networks in which at least 30% of the members are eligible. The second set of estimates of the Table are the reduced-form regression in treated village, regressing ineligible log-food consumption on potential rather than current log-transfer. The effects are positive and significant. The bottom part of the Table shows the results from our validation exercise: ineligible households' consumption in control villages is not a function of the potential transfer (i.e. of the number of potentially eligible children in the networks). Thus, the effects we estimated in treatment villages is caused by the program transfers, rather than by a comparison of networks with different consumption level irrespective of the program.

If connected eligible households share their transfers with their ineligible relatives, then we would expect connected eligible households to consume a smaller fraction of their transfer if they are related to some ineligibles, compared with connected eligibles who have no ineligible relatives.[8] We test this hypothesis by estimating treatment effects on consumption for connected eligible households with and without ineligible relatives. We then compare these treatment effects with the transfer size, computing the share of transfer consumed for these two groups separately. As a further consistency check, we compute consumption for their ineligible relatives as a fraction of the total network transfer. We present the results in Table 9. We estimate these effects first for food consumption only and then for all non-durable consumption, because, unlike the ineligibles, eligible households increase both food and non-food consumption in response to the program. We find that eligible connected households increase their food and total consumption by 28 and 34 pesos per adult per month out of a transfer of 64 pesos per adult, if

[8]The effect of Progresa on consumption depends on the transfer size, the presence of ineligible households in the network and the Pareto weights. If the latter are proxied by relative wealth in the network, eligible households in an all-eligible family network should share a smaller part of their transfer for two reasons: first, all network members are receiving the transfer; second, the Pareto weights of eligible households are on average higher in an all-eligible network.

they are related to ineligible households. Eligible households that have only eligible relatives, on the other hand, increase their consumption of 31 and 38 pesos respectively, out of a transfer of 55 pesos. Thus, connected eligible households consume a larger share of their transfer - 54% versus 69% - if they have ineligible relatives with whom they supposedly share their extra income. These latter households increase their consumption by 13 cents for each dollar transferred to their eligible relatives.

****************************** Comment about this table: we have to test whether the ATE shares are significanly different for households with and without ineligible relatives. we also should add the ATE share for isolated eligibles. Problem: they consume a tiny part of the transfer. We need to talk about it, as if we start showing this stuff then we have to explain what the isolated eligibles do with the progresa cash (in part they buy chickens but not enough to reconcile the small effect on consumption).

****

5.3 The effect on loans and gifts for ineligible households

Table 10 shows the indirect effect of Progresa of transfers and loans to connected and isolated households. As before, these effects are positive and significant for connected households and positive but smaller in magnitude and not significant for isolated households. Although the difference in the ITEs for connected and isolated households is never significant at the conventional levels (the ΔITEs for loans and transfers estimated by Tobit have t-tests of about 1.6), the pattern of results suggests that connected households use these additional transfers and loans to increase their consumption.

Since the quality of the loan and transfer data is not very good, the point estimates are probably not reliable. For example, the magnitude of estimates for loans changes according to the assumption one makes about outliers, since there are only a handful of households who receive transfers or loans. In the estimates presented above we trimmed the top 1% of positive values. When we estimated the ITE on loans without dropping the top percentile, the estimated effect becomes 21.4 pesos (with a standard error of 10.7), twice as large as the one presented above and accounting for the entire change in food consumption observed. Moreover, unlike for consumption, we also use data from November 1998 to increase the number of non-zero observations, as only XX% of households received either loans or transfers. Nevertheless, the treatment effects for loans and transfers are broadly consistent with our model predictions.

An additional explanation for the observed increase in consumption is that the program may increase ineligible households' income through increase of local wages or through its demand shock. However, Angelucci and De Giorgi (2007) show that there are no indirect treatment effect on labor and goods income and that prices do not change differentially between treatment and control villages. Moreover, if Progresa had also income effects we would expect consumption to increase also for isolated ineligibles, which is not the case.

Table 11 presents the estimates of the ITEs on loans and transfers. Since only a small fraction of households report receiving any loans or transfers, we group ineligible households according to whether they have any or no eligible relatives. We find that the indirect treatment effect on loans is positive and significant only for connected households who are related to program recipients. This group is 4.4 percentage point more likely to receive loans than connected households who are unrelated to program recipients, and receives on average 32 extra pesos, and 29 pesos more than connected households who are unrelated to eligible households. At the same time related households with eligible relative receive fewer transfers. Nevertheless, the increase in loans more than offsets the decrease in transfers.

In sum, in this section we have established that extended families form insurance subgroups: connected households share risk within each others, but not with isolated households.

5.4 Is it really the family?

In this section we investigate whether groups defined over land ownership (60% own any land in our sample) or ethnicity (35% of natives in our sample) do achieve similar level of insurance as the family dynasties. We split the sample between land owners and landless (as in Townsend, 1994) further we investigate whether the native population is able to smooth consumption better than non-natives. Unfortunately in the data there is not a clear variable defining indigenous households. We resort to the information on spoken languages (detail in the appendix). In the particular case there is no evidence that landowners are able to smooth consumption better than landless. The same results apply to natives, i.e. no evidence of better insurance is found among the natives compared to the rest of the population. Further when we investigate whether in the pooled samples, performing the joint test, we can find evidence that those latter two groups achieve full-insurance more frequently than families we find consistent evidence that this is not the case at all. In fact, for indigenous versus not-indigenous we cannot reject the hypothesis of full insurance for the former in roughly 1% of the villages. If we focus on land ownership such share goes only up to 5%; while, as we have seen earlier, the same percentage is close to 15% for connected versus isolated households.

As further evidence of the importance of the family dynasty in smoothing consumption fluctuations, we repeated the Townsend's village regressions cutting the sample according to ethnicity (i.e. indigenous households) and land ownership. In Tables 15 and **??** we can overall confirm that the family plays a relatively more important role as a smoothing device. In fact in Table 15, where we separately test for full insurance for the two groups, we find that the number of villages for which we reject the null are not too different between indigenous and non-indigenous as well as for land owners versus landless. We cannot reject full insurance for connected, landed and landless, indigenous and non-indigenous households in about 60% of the villages (at the 10% level). However, it is also clear that the share of rejections is not different in any of the data cut but that based on family dynasties. A note of caution is needed here in the interpretation of the results: while we know that isolated households are roughly a fourth of connected, for the other two cuts of the data that proportion is fairly different. In the data roughly 60% of the households own any land; while 35% of the households are indigenous in our definition (i.e. head of households speaks a native language). This clearly changes the sizes of the samples quite dramatically, in the three comparisons. On mere sample size we have the highest likelihood of not rejecting the null of full insurance in the case of isolated households, those are fewer than any other group considered in this part of the analysis.

6 The degree of insurance for connected households

6.1 Mace-Townsend analysis

The model presented in Section 2 suggests the use of the standard Mace-Townsend testing procedure to analyze the degree of insurance achieved by the different households in the Mexican villages. The standard estimating equation would be equation (**??**), where the test of full insurance reduces to testing whether $\beta_1 = 1$ and $\beta_2 = 0$.[9]

[9] It is well known that the test on the first parameter has virtually no informational content, given that it is true by construction.

We proceed here by implementing a number of sequential tests: first, we check whether full insurance is achieved in the aggregate, as in Townsend (2004). Second, we test whether it is achieved for connected households in the aggregate, in some villages or some networks. The logic of this exercise is the following: first, we provide baseline results taking the village as the relevant unit of risk sharing has done in the literature. Second, we proceed by considering the extended family as the relevant unit as shown in the previous part of the analysis (Section 5). We found that isolated households do not share risk with connected and further they do not seem to share risk with other isolated villagers.

At village level, including isolated households then, we reject the hypothesis of full-insurance. In particular we find a statistically significant effect of household income on household consumption. The coefficient (β_2) on household income in equation (5) is .02 (s.e. .002) see Table 14. As in Townsend (1994) the magnitude of the effect is rather small (in fact smaller then in Townsend), i.e. with no insurance at all that parameter should be 1 as a fully autarkic model with no savings suggests. Since we are concerned about the possible error in the measurement of our income variable, we perform the same type of testing by IV using the second lag of total household income, as well as measures of weather shocks, both as a simple dummy if any shock hit the household as well as the most frequent natural shock, i.e. drought, as well as a measure of health status, i.e activities of daily living (ADL) as in Gertler and Gruber (2002). ADL index is constructed from information available in wave 3 to 6 of the data on the abality to perform the following daily activities: 1) play soccer, run, wash clothes; 2) work in the orchard or walk 5 kilometers; 3) carry weight of 10 kilograms for 500 meters; 4) pick a paper from the floor; 5) walk more than 2 kilometers; 6) bathe and dress without help. Responses to those questions range from: 1, yes; 2) yes with some difficulty; 3) not at all. The results in Table **??** are confirmed, not surprisingly in Table **??** for connected households only, which we know constitute the vast majority of the households on the villages.

Table 15 perform a similar exercise to the one above, while looking into connected households only. Notice that we do not perform the same analysis for isolated households separately since we do not know what their insurance network, if any, looks like. Further, we have already shown how isolated households do not share risk with other households in the village. We would then be not able to interpret any result for the isolated households in the context of our model.

The main message from the table is that we reject efficient risk sharing for both connected and isolated taken together, i.e. at the village level. However it is also apparent that the idiosyncratic component has a small economic effect on the households. In fact, the estimate of a transformation of our transaction cost, is rather small, a 10% fluctuation in one's income translate in only a .2% change in consumption. Our estimates suggest a very low elasticity of consumption to idiosyncratic shocks irrespective of the shock measure we use.

Although we reject Pareto efficiency at the aggregate level for connected households, we notice a substantial heterogeneity in the degree of insurance achieved by the different networks of extended families. In fact, we devote Section **??** to this issue.

7 Is consumption smoothing costlier for the isolated?

This section first documents that consumption is smoother for connected and isolated households, and then tests whether isolated households engage in costlier activities to smooth consumption.

7.1 Consumption and income smoothness

****

Comment about this table: I would 1) drop the last column; 2) regress log Vcons on BOTH the I dummy and log V income (interacted by I?); 3) rather than showing food and non-food I'd show food and total consumption (it is consistent with the previous table with the ITE and ATE consumption shares)

****

Table **??** compares the logs in longitudinal coefficients of variation in consumption and income for connected and isolated households. Both food and non-food consumption are significantly more volatile for isolated than for connected households, by 6 and 3%. However, income is also significantly smoother for connected households: its longitudinal coefficient of variation is 2% smaller for connected households. On the aggregate measure of consumption, i.e. the network consumption, we cannot detect any significant difference between connected and isolated households. If we drop single headed households, the coefficients become smaller and less precisely estimated. The coefficient of variation of food consumption - the most reliable measure - is nevertheless still significantly larger for the isolated, by 2%.

7.2 Investment and negative shocks

****

The paragraphs below are very nice and I think they make an important point. However, this stuff right now does not come directly out of our model. We should be careful to reference to this when we discuss the implications of the model with insurance and investment

****

While consumption is smoother for connected than for isolated households, it varies much less than income for both groups (the longitudinal coefficients of variations of income are almost twice as large ad the coefficients of variation of consumption for both groups). This is probably not surprising. Since households are poor, they are likely very risk averse, therefore they probably engage in very costly consumption-smoothing activities when hit by a negative shock. Thus, the welfare of two households with similarly smooth consumption may differ substantially if one of the two is forced to sell productive assets, take their children out of school and have them work whenever hit by a bad shock. Therefore, besides comparing consumption smoothness for connected and isolated households, it is important to establish how costly smoothing consumption is for these two groups.

If isolated households are excluded from informal risk-sharing in the village, they likely engage in costly activities to smooth consumption when hit by negative shocks. In particular, isolated households may forego investment opportunities to stabilize their current consumption. Thus, households excluded from risk-sharing may face a trade-off between consumption and investment that insured households do not face. We test whether isolated households respond to health shocks by depleting their assets and livestock, by reducing schooling and increasing child labor more than connected households.

The empirical evidence suggests that isolated households employ more costly means of smoothing consumption through income smoothing devices. When hit by a weather shock, isolated households react by depleting their stock of chickens and oxen. That is, when we regress the change in the stock of chickens on the change in the weather shock status, the shock coefficients are significantly different for connected and isolated households, as shown in Tables **??** to **??**. Chickens are the most commonly held animals. Besides a larger drop in radio ownership if the household head falls sick, we find no other differential change in asset ownership for the connected and the isolated. More interestingly, it appears that the labor supply and school enrollment of young children in isolated households is significantly affected by a negative shock to the household. The picture drawn in Tables **??** and **??** is quite dramatic: secondary school aged children, between the ages of 11 and 16, respond to health shocks to the head and to weather shocks by withdrawing from school and supplying their labor services. For example, young children whose dad is sick in the last month work more often and with higher likelihood than their peers whose household is embedded in a family network. In particular, the effect on schooling is concentrated among boys as their are closer substitute to the sick father.

As shown in Table **??** isolated households react to the head of household falling sick by reducing enrollment of their children of secondary school age, i.e. 11 to 16 years old. The effect is concentrated among boys, closer substitute to their fathers in the labor market. On the contrary, we do not find differential response between isolated and connected households as far as absenteeism is concerned. It appears that the reaction is concentrated on the extensive rather than the intensive margin. However it is also true that boys from both households type are more frequently absent from school. Withdrawing children from school in the face of a shock to the head of the household can clearly contribute to smoothing income and therefore consumption, however it is an incredibly expensive technology. By reducing and hampering human capital accumulation, such behavior is bound to perpetrate poverty. On the other hand it also tell us something about the degree of risk aversion of the type of households in this analysis. As shown in Chetty and Looney (2006), the welfare loss from consumption fluctuation is proportional to the variance of consumption by a factor given by the risk aversion parameter. Such that even very small fluctuations of consumption could cause significant welfare losses.

7.3 Investment and positive shocks

8 Conclusion

TBC

References

[1] Albarran, Pedro and Orazio Attanasio (2002), "Do Public Transfers Crowd out Private Transfers? Evidence from a Randomized Experiment in Mexico". WIDER discussion paper

2002/6.

[2] Joseph G. Altonji, Fumio Hayashi, Laurence J. Kotlikof, "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," The American Economic Review, Vol. 82, No. 5 (Dec., 1992), pp. 1177-1198.

[3] Angelucci M. and De Giorgi, G., (2008), "Indirect Effects of an Aid Program: How do Cash Injections Affect Non-eligibles' Consumption, forthcoming the American Economic Review.

[4] Angelucci M., De Giorgi, G., Rangel, M.A., and Rasul, I. (2007), "Extended Family Networks in Rural Mexico: A Descriptive Analysis", forthcoming in CESifo Conference Volume on Institutions and Development, edited by Timothy Besley and Raji Jayaraman, Massachusetts: MIT Press.

[5] Attanasio, O., (2000), "Consumption", Handbook of Macroeconomics, Volume 1B, Taylor, J. and M. Woodford , eds., Amsterdam: Elsevier Science.

[6] Banerjee, A., (2004), "Inequality and Investment", mimeo, MIT.

[7] Banerjee, A., Duflo, E. and Munshi, K., (2003), "The (Mis)Allocation of Capital", Journal of the European Economic Association, 1(2-3): 484-494.

[8] Berhman, J., and Todd, P. (1999), "Randomness in the Experimental Sample of Progresa (Education, Health, and Nutrition Program)", International Food Policy Research Institute, Washington, D.C.

[9] Blundell, R., Costa-Dias, M., Meghir, C. and van Reenen, J., (2004), "Evaluating the Employment Impact of a Mandatory Job Search Program", Journal of the European Economic Association, 2(4): 569-606.

[10] Bobonis, G. and Finan, F. (2005), "Endogenous Social Interaction Effects in School Participation in Rural Mexico", mimeo, University of California Berkeley.

[11] Carroll, C. (2001), "A Theory of the Consumption Function, With and Without Liquidity Constraints," Journal of Economic Perspectives, 15(3), 23-46.

[12] Coate, S. and Ravallion, M. (1993), "Reciprocity without commitment Characterization and performance of informal insurance arrangements", Journal of Development Economics, 40: 1-24.

[13] Deaton, A. (1991), "Saving and Liquidity Constraints", Econometrica, 59(5): 1221-48.

[14] Deaton, A. (1992), "Saving and income smoothing in Cote dIvoire". J. Afr. Econ. 1 (1), 1 24 (March).

[15] Deaton, A. (1997), "The Analysis of Household Survey Data", Baltimore: Johns Hopkins University Press.

[16] Di Maro, V. (2004), "Evaluation of the Impact of Progresa on Nutrition: Theory, Econometric Methods and an Approach to Deriving Individual Welfare Findings from Household Data", mimeo, University College London.

[17] Chetty, Raj and Adam Looney, (2006), "Consumption Smoothing and the Welfare Consequences of Social Insurance in Developing Economies," Journal of Public Economics 90: 2351-2356, 2006

[18] Duflo, E. and Saez, E. (2002), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," unpublished working paper, MIT and University of California, Berkeley.

[19] Fafchamps, M. and Lund, S., (2003), "Risk-sharing Networks in Rural Philippines", Journal of Development Economics, 71(2): 261-87.

[20] Foster Andrew D. and Mark R. Rosenzweig, "Imperfect Commitment, Altruism, and the Family: Evidence from Transfer Behavior in Low- Income Rural Areas," Source: The Review of Economics and Statistics, Vol. 83, No. 3, (Aug., 2001), pp. 389-407

[21] Gertler, Paul and Jonathan Gruber, (2002), "Insuring Consumption against Illness,", the American Economic Review.

[22] Hoddinott, J., Skoufias, E., and Washburn, R. (2000) "The Impact of Progresa on Consumption: a Final Report". International Food Policy Research Institute, Washington, D.C.

[23] Kocherlakota, N.R. (1996), "Efficient bilateral risk sharing without commitment," Rev. Econ. Stud. 63 (4), 595 609 (October).

[24] Gertler, P.J., Martinez, Sebastian ,and Rubio-Codina, Marta (2006), "Investing Cash Transfers to Raise Long Term Living Standards", World Bank Policy Research Working Paper No. 3994.

[25] Lalive, R. and Cattaneo, A., (2005), "Social Interactions and Schooling Decisions", Mimeo, University of Zurich.

[26] Lim, Y. and Townsend, R. (1998), "General Equilibrium Models of Financial Systems: Theory and Measurement in Village Economies", Review of Economic Dynamics, 1(1): 59-118.

[27] Ligon, E., Thomas, J.P., and Worrall, T. (2002), "Informal insurance arrangemens with limited commitment: theory and evidence from village economies," Review of Economic Studies 69(1):209-244.

[28] Mazzocco, M and Saini, S. (2006),"Testing Efficient Risk Sharing with Heterogeneous Risk Preferences: Semi-parametric Tests with an Application to Village Economies," mimeo, UCLA.

[29] Miguel, E. and Kremer, M. (2004),"Worms: identifying impacts on education and health in the presence of treatment externalities," Econometrica, 72(1): 159-217.

[30] Munshi, K. and Rosenzweig, M., (2005),"Why is Mobility in India so Low? Social Insurance, Inequality, and Growth", CID Working Paper No. 121, Harvard.

[31] Parker, S. and Skoufias, E., (2000), "Final Report: The Impact of PROGRESA on Work, Leisure, and Time Allocation", International Food Policy Research Institute, Washington.

[32] Rosenzweig, M. (1988a), "Risk, Private Information, and the Family," American Economic Review, 78(2): 245-50.

[33] Rosenzweig, M. (1988b), "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries", The Economic Journal, 98(393): 1148-70.

[34] Rosenzweig, M.R., Stark, O., 1989. Consumption smoothing, migration, and marriage: evidence from rural India. J. Polit. Econ. 97 (4), 905926 (August).

[35] Rosenzweig, M.R., Wolpin, K.I., 1993. Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: investments in bullocks in India. J. Polit. Econ. 101 (2), 223244.

[36] Sacerdote, B. (2001), "Peer Effects with Random Assignment: Results for Dartmouth Roommates," Quarterly Journal of Economics, 116: 681-704.

[37] Schultz, P., (2001), "School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program", forthcoming in Journal of Development Economics.

[38] Skoufias, E., Davis, B. and Behrman, J., (1999a), "Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico", International Food Policy Research Institute, Washington, D.C.

[39] Skoufias, E., Davis, B. and de la Vega, S., (1999b), "An Addendum to the Final Report: An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico. Targeting the Poor in Mexico: Evaluation of the Selection of Beneficiary Households into PROGRESA", International Food Policy Research Institute, Washington, D.C.

[40] Schulhofer-Wohl, S. (2006), "Heterogeneity, Risk Sharing and the Welfare Costs of Risk," mimeo, U Chicago.

[41] Townsend, R., (1994), "Risk and Insurance in Village India", Econometrica, 62(3): 539-591.

[42] Townsend, R., (1995a), "Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies", The Journal of Economic Perspectives, 9(3): 83-102.

[43] Townsend, R., (1995b), "Financial Systems in Northern Thai Villages," The Quarterly Journal of Economics, 110(4): 1011-46.

[44] Udry, C. (1994), "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria," Review of Economic Studies, 61(3): 495-526.

[45] Udry, C. (1995), "Risk and Saving in Northern Nigeria", American Economic Review, 85(5): 1287-1300.

Table 1: Extended Family Network Descriptives

	Network Size	Size/# Households	Distance	Degree	Diameter
Mean	7.76	.167	2.12	2.07	2.45
SD	(9.65)	(.149)	(.303)	(1.20)	(1.18)

One observation per family network. 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 22549 couple headed households, 17029 (75.5%) of them are within family networks with at least two households. 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.

Table 2: Demographic and economic characteristics of connected and isolated households

	Couple-headed households					All households				
	Connected		Isolated			Connected		Isolated		
Variable	Mean	SD	Mean	SD	p-value	Mean	SD	Mean	SD	p-value
Household size:										
age 07	1.3	(1.26)	1.25	(1.28)	[0.059]*	1.24	(1.26)	0.97	(1.22)	[0.000]***
age 814	1.14	(1.21)	1.15	(1.22)	[0.817]	1.1	(1.2)	0.93	(1.16)	[0.000]***
age 1518	0.54	(0.78)	0.55	(0.78)	[0.610]	0.53	(0.77)	0.48	(0.75)	[0.000]***
age 1921	0.27	(0.54)	0.25	(0.52)	[0.062]*	0.27	(0.54)	0.23	(0.5)	[0.000]***
age 22+	2.38	(0.97)	2.4	(0.93)	[0.333]	2.33	(1)	2.19	(1.01)	[0.000]***
Household head:										
age	44.69	(15.11)	47.18	(15.84)	[0.000]***	45.55	(15.4)	51.62	(17.22)	[0.000]***
literacy	0.73	(0.44)	0.67	(0.47)	[0.000]***	0.71	(0.45)	0.59	(0.49)	[0.000]***
sex	0.99	(0.1)	0.99	(0.1)	[0.654]	0.93	(0.25)	0.77	(0.42)	[0.000]***
ethnicity	0.34	(0.48)	0.38	(0.49)	[0.172]	0.34	(0.48)	0.35	(0.48)	[0.584]
Couple headed	1	(0)	1	(0)	[.]	0.9	(0.3)	0.65	(0.48)	[0.000]***
Temp. mig.:										
US (dummy)	0.055	(0.229)	0.031	(0.174)	[0.001]***	0.058	(0.234)	0.038	(0.192)	[0.000]***
MX (dummy)	0.135	(0.342)	0.113	(0.317)	[0.200]	0.137	(0.344)	0.124	(0.329)	[0.014]**
Current mig.:										
US (dummy)	0.017	(0.128)	0.018	(0.133)	[0.601]	0.017	(0.128)	0.017	(0.131)	[0.741]
MX (dummy)	0.016	(0.124)	0.016	(0.124)	[0.988]	0.016	(0.125)	0.015	(0.122)	[0.719]
Land:										
Irrigated (dummy)	0.04	(0.2)	0.06	(0.23)	[0.075]*	0.04	(0.2)	0.05	(0.22)	[0.111]
Other (dummy)	0.6	(0.49)	0.57	(0.5)	[0.031]**	0.6	(0.49)	0.55	(0.5)	[0.000]***
Irrigated size	0.12	(1.01)	0.14	(0.75)	[0.292]	0.11	(0.97)	0.13	(0.72)	[0.399]
Other size	2.18	(4.2)	2.21	(4.91)	[0.766]	2.14	(4.11)	2.09	(4.57)	[0.558]
horse	0.4	(1.04)	0.39	(0.99)	[0.544]	0.39	(1.02)	0.34	(0.96)	[0.012]**
donkey	0.4	(1.16)	0.39	(0.87)	[0.570]	0.39	(1.13)	0.35	(0.82)	[0.007]***
ox	0.12	(0.74)	0.12	(1.01)	[0.789]	0.11	(0.73)	0.11	(0.88)	[0.688]
goat	1.58	(5.93)	1.4	(5.09)	[0.168]	1.6	(5.96)	1.42	(5.48)	[0.107]
cow	1.16	(3.76)	1.1	(4.12)	[0.520]	1.13	(3.7)	0.99	(3.7)	[0.078]*
chicken	7.01	(8.16)	7.98	(8.81)	[0.000]***	6.9	(8.12)	7.4	(8.55)	[0.003]***
pig	1.2	(2.92)	1.18	(3.11)	[0.782]	1.18	(2.95)	1.08	(2.84)	[0.103]
rabbit	0.22	(2.21)	0.25	(2.5)	[0.553]	0.22	(2.19)	0.22	(2.3)	[0.892]
blender	0.35	(0.48)	0.34	(0.47)	[0.306]	0.34	(0.48)	0.32	(0.47)	[0.084]*
fridge	0.16	(0.36)	0.16	(0.37)	[0.591]	0.15	(0.36)	0.16	(0.36)	[0.624]
stove	0.32	(0.47)	0.29	(0.45)	[0.039]**	0.32	(0.46)	0.29	(0.45)	[0.056]*
heater	0.03	(0.17)	0.03	(0.18)	[0.433]	0.03	(0.17)	0.03	(0.17)	[0.853]
radio	0.65	(0.48)	0.64	(0.48)	[0.403]	0.65	(0.48)	0.61	(0.49)	[0.001]***
CD player	0.05	(0.22)	0.06	(0.23)	[0.287]	0.05	(0.22)	0.05	(0.22)	[0.512]
TV	0.49	(0.5)	0.46	(0.5)	[0.015]**	0.48	(0.5)	0.43	(0.49)	[0.000]***
VHR	0.04	(0.19)	0.04	(0.2)	[0.388]	0.04	(0.19)	0.04	(0.18)	[0.501]
washer	0.05	(0.22)	0.05	(0.22)	[0.459]	0.05	(0.21)	0.05	(0.22)	[0.493]
fan	0.08	(0.27)	0.09	(0.28)	[0.260]	0.08	(0.27)	0.08	(0.28)	[0.383]
Car	0.02	(0.15)	0.03	(0.16)	[0.324]	0.02	(0.15)	0.02	(0.15)	[0.765]
truck	0.08	(0.28)	0.08	(0.27)	[0.464]	0.08	(0.27)	0.07	(0.25)	[0.005]***
Wealth index	728.43	(141.48)	728.9	(141.35)	[0.912]	732.68	(140.71)	747.78	(139.34)	[0.000]***
Income	272.79	(352.64)	275.99	(325.17)	[0.693]	279.49	(367.63)	291.11	(402.84)	[0.151]
Non-food exp	67.59	(85.43)	67.13	(87.98)	[0.830]	68.84	(91.02)	73.34	(119.4)	[0.047]**
Food exp (1)	153.65	(103.72)	155.69	(111.23)	[0.448]	160.36	(117.11)	181.89	(156.14)	[0.000]***
Food exp (2)	153.05	(127.35)	152.27	(136.44)	[0.829]	159.85	(141.9)	177.38	(187.8)	[0.000]***
Weekly child labor:										
Dummy age 8-14	0.12	(0.32)	0.13	(0.34)	[0.279]	0.12	(0.32)	0.13	(0.34)	[0.095]*
Dummy age 15-18	0.48	(0.5)	0.48	(0.5)	[0.972]	0.49	(0.5)	0.51	(0.5)	[0.234]
Hours age 8-14	2.63	(9.03)	2.66	(8.82)	[0.919]	2.65	(9.08)	2.87	(9.3)	[0.283]
Days 8-14	0.39	(1.26)	0.4	(1.25)	[0.784]	0.39	(1.26)	0.43	(1.3)	[0.239]
Hours age 15-18	17.73	(21.82)	17.46	(21.5)	[0.695]	18.07	(21.88)	18.69	(22.03)	[0.319]
Days 15-18	2.2	(2.58)	2.2	(2.58)	[0.943]	2.24	(2.59)	2.34	(2.62)	[0.172]

P-values of differences from standard errors clustered at the village level. September 1997 data, except March 1998 expenditures. Consumption and income per adult equivalent and at November 1998 pesos.

Table 3: Correlation of natural disasters in a pair of households

	Same network	Same village
	(1)	(2)
Drought	0.29	0.27
SE	0.01	0.02
Flood	0.21	0.19
SE	0.02	0.05
Frost	0.27	0.24
SE	0.02	0.05
Fire	0.08	0.14
SE	0.04	0.15
Pest	0.17	0.19
SE	0.02	0.05

Note: correlation computed for two randomly selected households in a network of degree one or any two households in a village. Results based on 100 replications. Data on shocks are available in wave 3 to 6.

Table 4: Coefficients of Variation (CV=Var/Mean) of consumption and income for connected and isolated households

	All Households		Couple-Headed	
	Connected	Isolated	Connected	Isolated
	Food Consumption			
	Household level statistics			
Mean	192.18	196.13	184.54	176.11
sd	130.79	88.92	114.12	72.01
Cross-sectional CV	0.44	0.56	0.43	0.48
sd	0.30	0.30	0.29	0.27
Longitudinal CV	0.37	0.40	0.36	0.37
sd	0.20	0.23	0.19	0.21
	Aggregate Village Consumption			
Longitudinal CV: hhold/village	2.26	2.46	2.31	2.31
sd	1.72	1.98	1.73	1.73
	Aggregate Network Consumption			
Longitudinal CV: hhold/network	1.54	1.75	1.52	1.64
sd	1.16	1.40	1.16	1.34
	Non-Food, non-Durable Consumption			
Mean	74.78	72.35	73.93	69.87
sd	59.41	45.25	61.9	50.92
Cross sectional CV	0.66	0.84	0.65	0.75
sd	0.36	0.34	0.35	0.33
Longitudinal CV	0.70	0.72	0.69	0.70
sd	0.30	0.32	0.30	0.30
	Total Income			
Mean	223.53	220.78	221.02	217.81
sd	148.13	119.54	150.4	132.95
Cross sectional CV	0.72	0.86	0.39	0.36
sd	0.40	0.35	0.69	0.71
Longitudinal CV	0.70	0.73	0.69	0.71
sd	0.34	0.36	0.34	0.34

Note:

Monthly peso value per adult equivalent (10 pesos≈1 USD). One observation per network. Extreme values trimmed in the computation above. For isolated households the cross-sectional coefficient of variation is computed assuming that all isolated households in a village are in the same network. The longitudinal coefficient of variation is computed for all households and then one observation per network is used for the descriptive statistics.

Table 5: Changes in income, consumption, investment, and savings for connected and isolated households - Difference-in-difference estimation between 1997 and 2003 - couple-headed households only.

	Labor	Non-durable consumption:		Durable consumption. % owning:				
	income	Food	Non-food	Fridge	Stove	Heater	Radio	TV
Constant	494.19	277.1	105.12	0.045	0.362	0.036	0.544	0.471
	[16.749]***	[7.327]***	[4.741]***	[0.021]**	[0.037]***	[0.008]***	[0.020]***	[0.027]***
I	11.85	2.21	3.23	0.026	0.000	0.007	0.006	-0.010
	[6.485]*	[2.584]	[1.609]**	[0.009]***	[0.011]	[0.004]*	[0.010]	[0.012]
2003	30.01	-11.64	56.88	0.174	0.081	0.014	-0.052	0.157
	[5.361]***	[2.324]***	[2.055]***	[0.008]***	[0.007]***	[0.004]***	[0.010]***	[0.008]***
I*2003	-15.11	-4.82	-2.98	-0.024	-0.020	-0.016	-0.001	-0.024
	[8.222]*	[3.136]	[2.711]	[0.010]**	[0.010]**	[0.005]***	[0.015]	[0.011]**
Obs.	35093	35433	35416	35642	35645	35598	35632	35641

	Education: % with at least		Savings/Investment: 1. Stock of animals					2. % owning machinery	
	Primary	Secondary	Cow	Ox	Horse/donkey	Chicken	Goat/pig	Car	Tractor
Constant	0.444	0.086	0.049	0.003	0.124	1.574	0.511	0.025	-0.002
	[0.022]***	[0.012]***	[0.027]*	[0.005]	[0.015]***	[0.095]***	[0.051]***	[0.010]**	[0.003]
I	-0.017	0.015	-0.019	-0.001	-0.005	0.135	-0.016	0.013	-0.001
	[0.007]**	[0.004]***	[0.015]	[0.003]	[0.006]	[0.044]***	[0.029]	[0.004]***	[0.001]
2003	0.053	0.045	-0.115	-0.004	-0.076	-0.633	-0.242	0.130	0.004
	[0.004]***	[0.003]***	[0.009]***	[0.002]**	[0.005]***	[0.033]***	[0.017]***	[0.008]***	[0.001]***
I*2003	-0.002	-0.009	0.033	-0.004	-0.004	-0.112	0.024	-0.021	-0.001
	[0.005]	[0.004]**	[0.014]**	[0.003]	[0.006]	[0.050]**	[0.027]	[0.009]**	[0.002]
Obs.	35661	35661	35325	35390	35244	35116	35214	35656	35652

Note: double difference in consumption (durable and non-durable) and machinery between Nov. 1998 and Nov. 2003, as data not available in 1997 (or available but potentially unreliable, as in 1997 consumption). Standard errors clustered at the village level. Differential trends allowed in two ways: 1) different trends for isolated and connected; 2) different trends by region, household and village poverty, and observable characteristics with different means for connected and isolated (household demographics and literacy, land ownership). Couple-headed households only. Including also single-headed households does not change the results.

Table 6: Extended families and the indirect effect of Progresa on food consumption

	(1)	(2)	(3)	(4)	(5)
	Pre-program	1999			
		Diff-in-diff		Cross Section	
		All households			
ITE^K	-10.65	28.45	14.20	17.39	10.44
	[7.52]	[12.60]**	[9.18]	[8.90]**	[6.09]*
Obs	3613	6366	6335	6366	6335
ITE^I	8.71	-13.94	-9.88	-6.22	2.46
	[12.41]	[17.07]	[15.55]	[13.91]	[11.19]
Obs	1227	2103	2094	2103	2094
ΔITE	-19.36	42.39	4.32	23.60	7.98
	[12.27]	[18.86]**	[15.19]	[15.62]	[12.11]
		Couple-headed only			
ITE^K	-8.87	26.54	18.58	17.04	10.87
	[7.13]	[11.20]**	[8.90]**	[8.42]**	[6.40]*
Obs	3138	5521	5503	5521	5503
ITE^I	-11.44	-14.73	1.29	-28.62	-6.69
	[10.92]	[20.14]	[16.56]	[17.98]	[12.48]
Obs	828	1469	1464	1469	1464
ΔITE	2.57	41.27	17.29	45.66	17.56
	[10.97]	[20.80]**	[16.47]	[19.08]***	[13.28]

Difference-in-difference OLS estimates, controlling for 1997 household characteristics, state and time effects. Standard errors clustered at the village level. Columns (2) and (4) drop all consumption data bigger than 10,000, while columns (3) and (5) trim the top and bottom 0.25%.

Table 7: The indirect effect of Progresa on food consumption for network members with and without eligible relatives

	(1)	(2)	(3)		(4)
	No eligible relatives	Up to 30% eligible	$>$ 30% eligible	Network size (n)	
		Cross section			
ITE^K	-15.18	0.20	19.28		
	[30.85]	[18.51]	[9.80]**		
$\text{ITE}^K_{>30\%}$-ITE^K	34.45	19.07	—		
	[31.82]	[28.80]	—		
Obs.		6366			
		Difference-in-difference			
ITE^K	21.26	-3.26	31.96	$\text{ITE}^K_{n<7}$	-7.04
	[39.20]	[25.56]	[13.14]**		[16.92]
$\text{ITE}^K_{>30\%}$-ITE^K	10.69	35.21	—	$\text{ITE}^K_{n\geq 7}$	42.11
	[38.71]	[28.17]	—		[14.37]***
Obs.		9979		Obs.	9979

Difference-in-difference OLS estimates, controlling for 1997 household characteristics, state and time effects. The percentage of eligible is the fraction of eligible households in the network. Standard errors clustered at the village level.

Table 8: Ineligible log-food consumption as a function of log-transfer per network member

		Treatment villages	
	OLS	IV	IV p>.30
ln(actual transfer)	0.078	0.133	0.205
	[0.025]***	[0.058]**	[0.097]**
Obs.	3353	3353	2409
		OLS	OLS p>.30
ln(potential transfer)		0.091	0.143
		[0.041]**	[0.067]**
Obs.		3353	2409
		Control villages	
		OLS	OLS p>.30
ln(potential transfer)		-0.006	-0.169
		[0.050]	[0.111]
Obs.		1230	829

OLS and IV estimates using log-potential transfer as IV for log-actual transfer per network adult, controlling for 1997 household and network characteristics, state and time effects. p is the fraction of eligible households in the network. The consumption elasticity of the transfer is 0.37 sd 0.08 (0.38 sd 0.08) for all treated (for p>.30) households. Standard errors clustered at the village level.

Table 9: Treatment effects on monthly consumption per adult equivalent for connected eligible households, 1999.

	Food consumption	All consumption
	Networks with both eligibles and ineligibles	
ATE	28.104	34.355
	[5.860]***	[7.646]***
Obs	10883	10883
ITE per eligible	8.512	8.618
	[3.517]**	[4.768]*
Obs	5478	5472
Transfer per eligible	63.93	63.93
ATE share	**0.44**	**0.54**
ITE share	0.13	0.13
	Networks with eligibles only	
ATE	31.226	38.569
	[7.421]***	[10.358]***
Obs	5627	5627
Transfer per eligible	55.5	55.5
ATE share	**0.56**	**0.69**

OLS first-difference estimates of consumption using May 1999, and November 1999 and subtracting March 1998 consumption, controlling for 1997 household characteristics, state and time effects. Standard errors clustered at the village level.

Table 10: Extended families and the indirect effect of Progresa on loans and transfers

	Informal loans			Transfers		
	Likelihood	Amount		Likelihood	Amount	
	Probit	OLS	Tobit	Probit	OLS	Tobit
	All ineligible households					
ITE connected	0.014	8.569	4.58	0.007	2.51	1.12
	[0.007]**	[4.395]*	[2.65]*	[0.005]	[1.29]*	[0.46]***
ITE isolated	-0.001	0.806	0.60	0.004	1.87	0.22
	[0.010]	[7.223]	[4.52]	[0.008]	[2.62]	[0.70]
ΔITE	0.015	7.76	2.98	0.003	0.63	0.90
	[0.011]	[8.32]	[3.97]	[0.005]	[2.41]	[0.86]
Obs.	9105	9097	9097	9012	9163	9163
	Couple-headed ineligible households					
ITE connected	0.012	7.897	5.91	0.006	2.068	1.12
	[0.007]*	[4.619]*	[3.14]*	[0.004]*	[1.247]*	[0.49]**
ITE isolated	-0.003	-7.163	-2.32	-0.002	2.203	-0.13
	[0.011]	[9.365]	[6.15]	[0.006]	[1.844]	[0.88]
ΔITE	0.015	15.06	8.23	0.008	-0.13	1.25
	[0.011]	[9.40]	[6.85]	[0.006]	[1.98]	[1.01]
Obs.	7541	7530	7530	7557	7562	7562

OLS estimates pooling the three post-program waves for November 1998, May 1999, and November 1999, controlling for 1997 household characteristics, state and time effects. The estimated coefficients without adding any covariates are similar in magnitude and significance to the ones reported above. Standard errors clustered at the village level.

Table 11: The indirect effect of Progresa on loans and transfers for connected households with and without related program recipients

	Informal loans			Transfers		
	Likelihood	Amount		Likelihood	Amount	
	Probit	OLS	Tobit	Probit	OLS	Tobit
	Connected ineligible households					
ITE eligible	0.050	32.81	31.62	0.006	-3.688	-0.59
relatives	[0.029]**	[10.61]***	[8.27]***	[0.006]	[5.966]	[0.99]
ITE ineligble	0.006	3.95	2.22	0.008	3.025	1.12
relatives	[0.006]	[4.31]	[2.63]	[0.005]	[1.253]**	[0.43]***
ΔITE	0.044	28.86	29.40	-0.002	-6.71	-1.71
	[0.024]*	[11.01]***	[14.41]**	[0.001]*	[6.07]	[1.04]*
Obs.	6754	6775	6775	6724	6849	6849

OLS estimates pooling the three post-program waves for November 1998, May 1999, and November 1999, controlling for 1997 household characteristics, state and time effects. The estimated coefficients without adding any covariates are similar in magnitude and significance to the ones reported above. Standard errors clustered at the village level.

Table 12: Treatment effects on 1999 food consumption for ineligible households by network size, land ownership, and ethnicity.

Network size		Land ownership		Ethnicity	
ITE^K n<7	-7.04	ITE land	18.23	ITE Indigenous	30.98
	[16.92]		[12.40]		[24.43]
ITE^K n$\geq$7	42.11	ITE no land	13.72	ITE Hispanic	14.05
	[14.37]***		[15.25]		[11.95]
DeltaITE	49.15	DeltaITE	4.51	DeltaITE	16.92
	[19.74]**		[17.05]		[26.82]
Obs	13468	Obs	13468	Obs	13468

Difference-in-difference OLS estimates. Standard errors clustered at the village level.

Table 13: treatment effects on 1999 food consumption for ineligible households by land ownership, and ethnicity over and above effect of family network.

Land ownership		Ethnicity	
ITE$^K_{L1}$	28.69	ITE$^K_{I1}$	42.16
	[13.84]**		[26.41]
ITE$^K_{L0}$	23.38	ITE$^K_{I0}$	24.16
	[18.35]		[13.48]*
ITE$^K_{L1}$-ITE$^K_{L0}$	5.30	ITE$^K_{I1}$-ITE$^K_{I0}$	18.23
	[20.90]		[12.40]
ITE$^I_{L1}$	-17.27	ITE$^I_{I1}$	-2.42
	[21.73]		[36.62]
ITE$^I_{L0}$	-5.04	ITE$^I_{I0}$	-16.29
	[24.34]		[18.59]
ITE$^I_{L1}$-ITE$^I_{L0}$	-12.23	ITE$^I_{L1}$-ITE$^I_{L0}$	13.72
	[32.39]		[15.25]
Obs	13468	Obs	13468

Difference-in-difference OLS estimates. Standard errors clustered at the village level.

Table 14: Full insurance regressions by household type

	Specification 1	Specification 2	Specification 3	Specification 1	Specification 2	Specification 3
	ALL HOUSEHOLDS			COUPLE HEADED HOUSEHOLDS		
			Total Income			
LnC	0.97	0.98	0.99	0.92	0.93	0.97
se	[0.007]***	[0.017]***	[0.01]***	[0.011]***	[0.011]***	[0.001]***
Lny	0.02	0.02	0.02	0.02	0.02	0.02
se	[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.002]***
Obs.	72855	72855	72855	62321	62321	62321
			Total Income IV (Lag 2 Income)			
LnC	0.97	0.98	0.99	0.93	0.94	0.96
se	[0.012]***	[0.013]***	[0.019]***	[0.014]***	[0.015]***	[0.019]***
Lny	0.00	0.01	0.02	0.00	0.00	0.01
se	[0.018]	[0.018]	[0.019]	[0.018]	[0.019]	[0.02]
Obs.	52607	52607	52607	43262.00	43262.00	43262.00
			Any Shock			
LnC	0.97	0.98	1.00	0.93	0.94	0.98
se	[0.007]***	[0.007]***	[0.01]***	[0.011]***	[0.012]***	[0.01]***
Shock	0.03	0.03	0.03	0.03	0.03	0.03
se	[0.007]***	[0.007]***	[0.007]***	[0.007]***	[0.007]***	[0.007]
Obs.	73202	73202	73202	62602	62602	62602
			Drought			
LnC	0.98	0.98	1.00	0.93	0.94	0.98
se	[0.007]***	[0.008]***	[0.01]***	[0.014]***	[0.015]***	[0.01]***
Drought	0.04	0.04	0.04	0.04	0.04	0.04
se	[0.009]***	[0.009]***	[0.009]***	[0.008]***	[0.009]***	[0.009]***
Obs.	56997	56997	56997	48619	48619	48619
			Activities of Daily Living (ADL index)			
LnC	0.95	0.96	1.01	0.91	0.92	0.98
se	[0.01]***	[0.011]***	[0.013]***	[0.014]***	[0.014]***	[0.013]***
ADL	-0.06	-0.07	-0.07	-0.06	-0.07	-0.07
se	[0.017]***	[0.018]***	[0.019]***	[0.017]***	[0.017]***	[0.019]***
Obs.	34267	34267	34267	29358	29358	29358

Notes: Specifications stands for different definition of aggregate consumption. Specification 1: for connected aggregate consumption is computed as the sum of consumption in the network; for isolated is the sum of consumption of isolated households in the village. Specification 2: For connected aggregate consumption is computed as the sum of consumption in the network; for isolated is total village consumption. Specification 3: Aggregate consumption is total village consumption. SE clustered at the village level. Extreme values trimmed in the regressions Consumption equivalent smaller than 10000 pesos. ***, **, * significantly different from zero at 1, 5, 10%. ADL index takes values between 0 and 1, 0 if no difficulties and 1 if unable to perform any of the activities.

Table 15: Full-insurance regressions in Extended Families $\Delta \ln c_{ht} = a + \alpha \Delta \ln C_t + \beta \Delta \ln y_{ht}$.

	ALL HOUSEHOLDS		COUPLE HEADED HOUSEHOLDS	
	Specification 1	Specification 3	Specification 1	Specification 3
Total Income				
LnC	0.97	0.96	0.93	0.96
se	[0.007]***	[0.012]***	[0.012]***	[0.013]***
Lny	0.02	0.02	0.02	0.02
se	[0.002]***	[0.002]***	[0.002]***	[0.002]***
Obs.	55965	55965	51055	51055
Total Income IV (Lag 2 Income)				
LnC	0.97	0.97	0.93	0.95
se	[0.0125]***	[0.022]***	[0.016]***	[0.023]***
Lny	0.00	0.01	0.00	0.02
se	[0.02]	[0.022]	[0.021]	[0.023]
Obs.	40502	40502	35552	35552
Any Shock				
LnC	0.97	0.97	0.93	0.97
se	[0.007]***	[0.011]***	[0.013]***	[0.012]***
Shock	0.03	0.03	0.03	0.03
se	[0.007]***	[0.007]***	[0.008]***	[0.008]***
Obs.	56214	56214	51278	51278
Drought				
LnC	0.97	0.98	0.93	0.97
se	[0.008]***	[0.012]***	[0.016]***	[0.013]***
Drought	0.03	0.03	0.04	0.04
se	[0.01]***	[0.01]***	[0.01]***	[0.01]***
Obs.	43687	43687	39800	39800
Activities of Daily Living (ADL index)				
LnC	0.94	0.98	0.91	0.97
se	[0.011]***	[0.015]***	[0.015]***	[0.164]***
ADL	-0.06	-0.06	-0.06	-0.07
se	[0.02]***	[0.022]***	[0.019]***	[0.02]***
Obs.	26374	26374	24098	24098

Notes: Specifications stands for different definition of aggregate consumption. Specification 1: for connected aggregate consumption is computed as the sum of consumption in the network. Specification 3: Aggregate consumption is total village consumption. SE clustered at the village level. Extreme values trimmed in the regressions. ***, **, * significantly different from zero at 1, 5, 10%. ADL index takes values between 0 and 1, 0 if no difficulties and 1 if unable to perform any of the activities.

Table 16: Tests on the variance of consumption and income

Log of CV:	Food	Non-Durables (Excl. Food)	Total Income	Aggregate Food
PANEL A: ALL HOUSEHOLDS				
Isolated	0.06	0.03	0.02	-0.03
SE	[0.010]***	[0.009]***	[0.012]**	[0.024]
Constant	-1.14	-0.46	-0.49	-1.36
SE	[0.007]***	[0.006]***	[0.009]***	[0.011]***
Observations	22200	22178	20169	2697
PANEL B: COUPLE HEADED HOUSEHOLDS				
Isolated	0.02	0.01	0.01	-0.04
SE	[0.012]*	[0.010]	[0.014]	[0.026]
Constant	-1.14	-0.47	-0.50	-1.38
SE	[0.007]***	[0.006]***	[0.010]***	[0.012]
Observations	18745	18740	17223	2635

Note: All households in a given network, isolated households are assumed to form a within village network. All variables refer to the monthly adult equivalent measures. Standard errors clustered at the village level. Data from the 5 available waves. The LHS variables represent the household specific time-series variances. Extreme values trimmed in the computation above. ***, **, * significant at 1, 5, 10%.

Table 17: Effect of Progresa and of household head illness on child labor, schooling, and livestock - first difference

	Child labor for 12-16 year olds			Schooling for 12-16 year olds			Stock of animals owned				
Effect on:	% working	At least one child works	Weekly workdays	Days lost	% enrolled	stock with complete secondary educ.	Cow	Ox	Horse/donkey	Chicken	Goat/pig
Positive shock: Progresa eligibility (randomized)											
Constant	0.188	0.314	0.940	2.167	-0.112	0.223	-0.005	0.005	0.014	-0.075	0.091
	[0.026]***	[0.033]***	[0.135]***	[0.584]***	[0.022]***	[0.054]***	[0.017]	[0.003]	[0.011]	[0.070]	[0.044]**
I	0.002	-0.003	-0.004	-0.453	0.013	-0.020	0.000	-0.001	-0.003	-0.091	0.001
	[0.009]	[0.012]	[0.049]	[0.236]*	[0.009]	[0.024]	[0.006]	[0.001]	[0.003]	[0.030]***	[0.016]
I*eligible	-0.006	-0.004	0.004	0.971	-0.029	0.003	0.004	0.002	0.003	0.064	-0.006
	[0.011]	[0.015]	[0.060]	[0.283]***	[0.011]***	[0.030]	[0.008]	[0.002]	[0.004]	[0.037]*	[0.021]
Eligible	-0.017	-0.020	-0.091	-0.787	0.032	0.011	0.008	-0.003	0.000	0.047	0.005
	[0.006]***	[0.008]**	[0.031]***	[0.142]***	[0.005]***	[0.014]	[0.004]*	[0.001]***	[0.003]	[0.022]**	[0.010]
Obs.	15236	15236	15236	16798	16798	5119	28849	28856	28668	28512	28647
Negative shock: houhsehold head illness											
Constant	0.037	0.046	0.245	1.440	-0.060	0.141	0.007	-0.002	0.032	0.100	0.072
	[0.014]***	[0.019]**	[0.076]***	[0.407]***	[0.015]***	[0.050]***	[0.021]	[0.003]	[0.009]***	[0.062]	[0.034]**
I	-0.005	-0.009	-0.025	-0.096	0.010	-0.017	-0.004	0.001	-0.002	-0.012	-0.009
	[0.004]	[0.005]*	[0.021]	[0.116]	[0.004]**	[0.013]	[0.006]	[0.001]	[0.003]	[0.018]	[0.010]
I*Ill	0.034	0.050	0.199	0.898	-0.038	-0.087	-0.011	0.002	-0.011	-0.012	0.048
	[0.019]*	[0.027]*	[0.102]*	[0.552]	[0.020]*	[0.051]*	[0.023]	[0.005]	[0.010]	[0.070]	[0.034]
Ill	0.018	0.021	0.080	0.374	-0.006	-0.002	-0.001	0.002	-0.001	0.095	-0.035
	[0.008]**	[0.011]*	[0.042]*	[0.264]	[0.009]	[0.023]	[0.011]	[0.002]	[0.005]	[0.032]***	[0.017]**
Obs.	21851	21851	21851	24026	24026	6884	31079	31148	30892	30558	30825

Note: double differences using the maximum number of available data waves 1) Sept. 1997 to Nov. 1999 data for effect of Progresa (using one wave per school year only for effect on stock with at least secondary education); 2) Nov 1998 to Nov 2000 for sickness of household head; 3) May 1999 to Nov 2000 for ADL. Data on animals owned available up to Nov. 1999. Standard errors clustered at the village level. Differential trends allowed in two ways: 1) different trends for isolated and connected; 2) different trends by region, household and village poverty, and observable characteristics with different means for connected and isolated (household demographics and literacy, land ownership). Couple-headed households only. ***, **, *: significant at 1, 5, 10%.

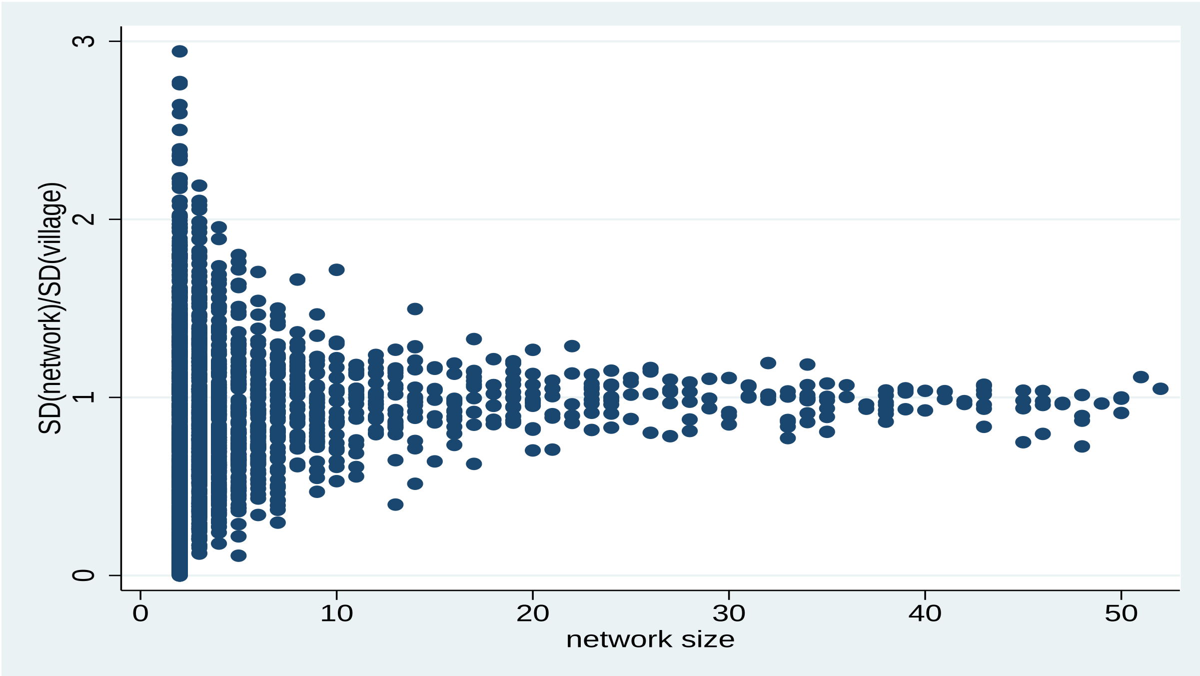
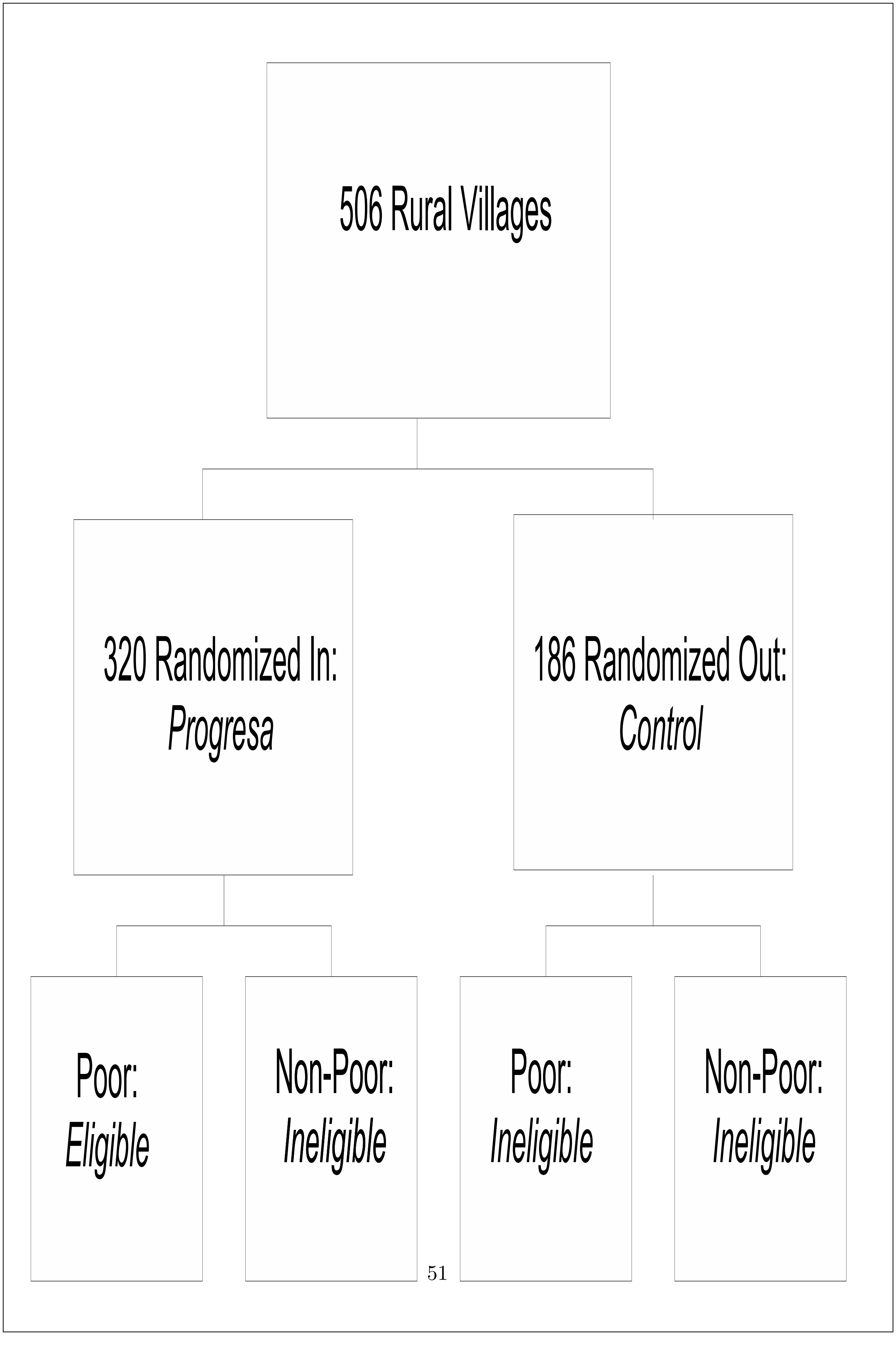
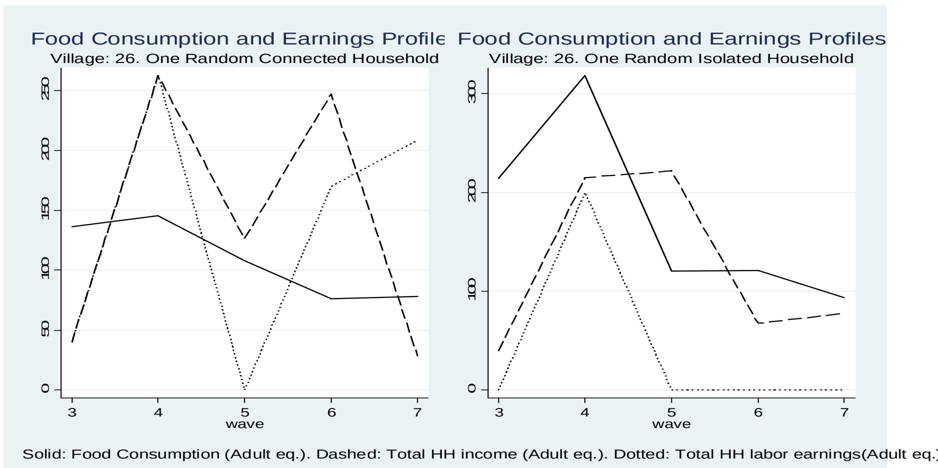
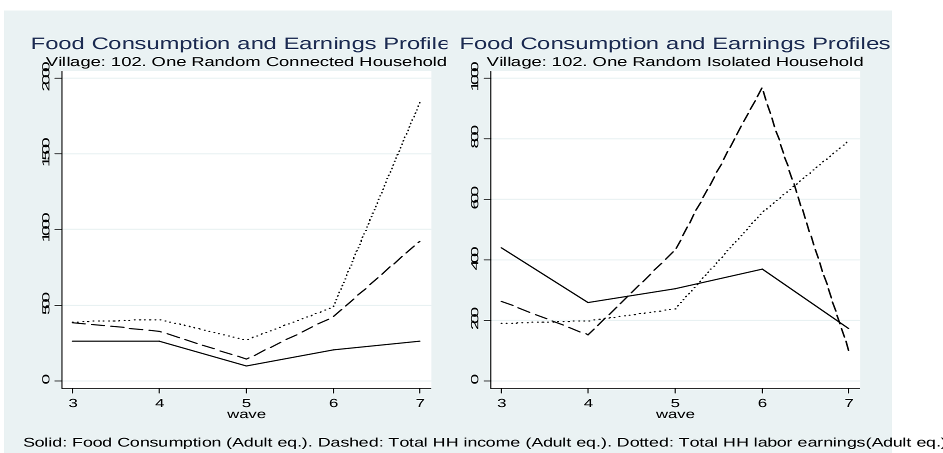


Figure 1: Ratio of network and village standard deviation of the wealth index


506 Rural Villages
320 Randomized In: *Progresa*
186 Randomized Out: *Control*
Poor: *Eligible*
Non-Poor: *Ineligible*
Poor: *Ineligible*
Non-Poor: *Ineligible*

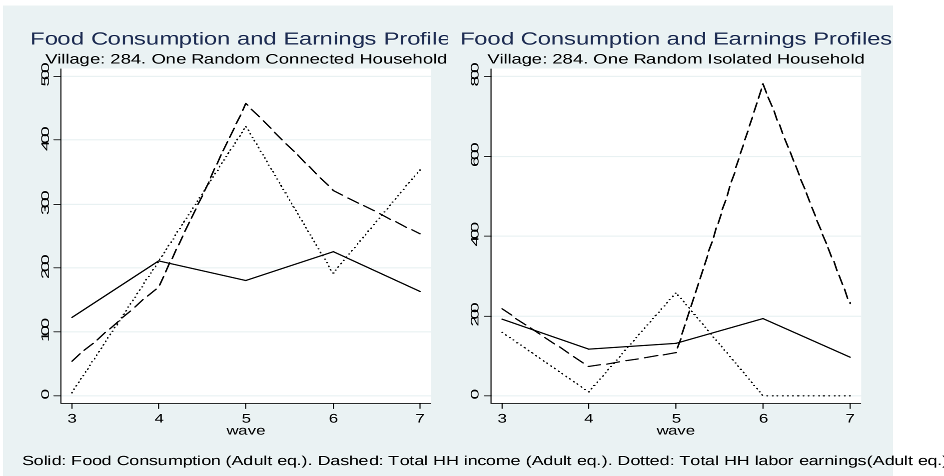


Note: This is for a village at the 25th percentile of the village size distribution (conditional on no missing in relevant info). CV(Consumption, connected)=.29; CV(Consumption, isolated)=.54. Village 26: is in the State of Guerrero

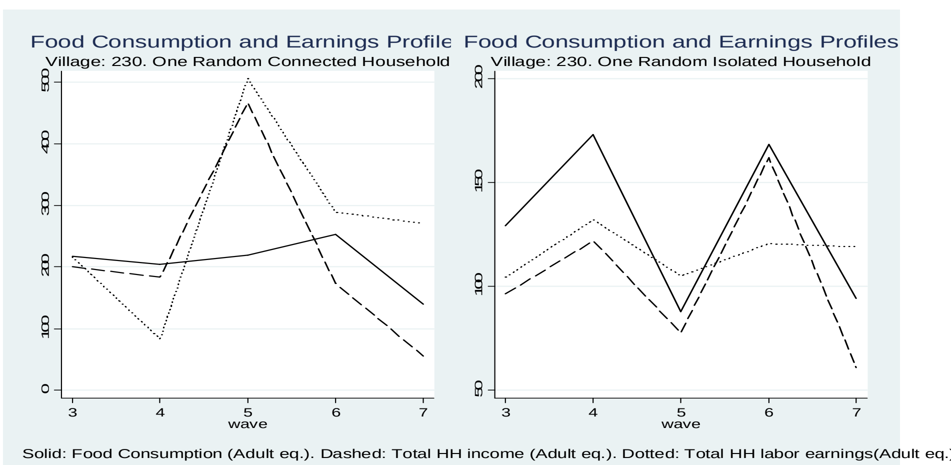


Note: This is for a village at the 50th percentile of the village size distribution (conditional on no missing in relevant info). CV(Consumption, connected)=.32; CV(Consumption, isolated)=.33. Village 102: is in the State of Hidalgo

Figure 2: Time profiles of food consumption and income



Note: This is for a village at the 50th percentile of the village size distribution (conditional on no missing in relevant info). CV(Consumption, connected)=.22; CV(Consumption, isolated)=.30.
Village 284: is in the State of Queretaro



Note: This is for a village at the 75th percentile of the village size distribution (conditional on no missing in relevant info). CV(Consumption, connected)=.20; CV(Consumption, isolated)=.31.
Village 230: is in the State of Puebla.

Figure 3: Time profiles of food consumption and income