

DISCUSSION PAPER SERIES

No. 3341

SOCIAL NETWORKS AND TECHNOLOGY ADOPTION IN NORTHERN MOZAMBIQUE

Oriana Bandiera and Imran Rasul

PUBLIC POLICY



Centre for Economic Policy Research

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP3341.asp

SOCIAL NETWORKS AND TECHNOLOGY ADOPTION IN NORTHERN MOZAMBIQUE

Oriana Bandiera, London School of Economics (LSE) and CEPR
Imran Rasul, London School of Economics (LSE)

Discussion Paper No. 3341
April 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Oriana Bandiera and Imran Rasul

April 2002

ABSTRACT

Social Networks and Technology Adoption in Northern Mozambique*

Despite their potentially strong impact on poverty, agricultural innovations are often adopted slowly. Using a unique household dataset on sunflower adoption in Mozambique, we analyse whether and how individual adoption decisions depend upon the choices of others in the same social networks. Since farmers anticipate that they will share information with others, we expect farmers to be more likely to adopt when they know many other adopters. Dynamic considerations, however, suggest that farmers who know many adopters might strategically delay adoption to free-ride on the information gathered by others. We present empirical evidence that shows that the relationship between the probability of adoption and the number of known adopters is shaped as an inverse-U. In line with information sharing, the network effect is stronger for farmers who report discussing agriculture with others. The data contains information that is needed to ameliorate the identification issues that commonly arise in this context. In particular, social networks are precisely identified, and in addition we can control for village heterogeneity and for endogenous group formation.

JEL Classification: O12 and O31

Keywords: information sharing, social networks and technology adoption

Oriana Bandiera
Department of Economics
London School of Economics
Houghton Street
LONDON
WC2A 2AE
Tel: (44 20) 7955 7519
Fax: (44 20) 7055 6951
Email: o.bandiera@lse.ac.uk

Imran Rasul
STICERD
London School of Economics
Houghton Street
LONDON
WC2A 2AE
Tel: (44 20) 7955 7058
Email: i.rasul@lse.ac.uk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=147704

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157538

* We would like to thank Tim Besley, Robin Burgess, Marcel Fafchamps, Raquel Fernandez, Andrew Foster, Markus Goldstein, Boyan Jovanovic, Magnus Lindelow, Ted Miguel, Rohini Pande, Debraj Ray, Giorgio Topa and seminar participants at New York University, LSE, NEUDC 2001 and CSAE, for useful comments.

Submitted 29 March 2002

1 Introduction

The adoption of new agricultural techniques is a key route out of poverty for many in the developing world. Yet, agricultural innovations have often been adopted slowly and some aspects of the adoption process are still poorly understood. Recent studies have shown that, both in developing and developed countries, social networks and peer effects are an important determinant of individual behaviour in a variety of settings. For instance, the literature in labour and public economics suggest that many individual decisions, as diverse as school attendance, drug use and welfare participation, are, both in theory and in practice, positively correlated with the behaviour of the social group the individual belongs to.¹ Macroeconomists have identified social learning and information spillovers as an important driving force in models of endogenous growth.² In the context of rural economies, it has been shown (Foster and Rosenzweig (1995), Conley and Udry (2000)) that farmers within a group learn from each other how to grow new crop varieties.³ Overall the evidence suggest that network effects are important for individual decisions, and that, in the particular context of agricultural innovations, farmers share information and learn from each other.

In this paper we analyse whether and how a farmer's decision to adopt a new technology depends upon the adoption decision of other farmers in his social group, which, unlike most of the existing literature, we are able to identify precisely. The analysis employs unique data on the adoption of sunflower for about two hundred farmers across nine villages, in the Zambezia province, Northern Mozambique. Data were collected in November-December 2000 together with the NGO *Movimondo*, who first introduced sunflower seeds in the area at the beginning of that year. Sunflower seeds were distributed free of charge as a part of a programme aimed at promoting food security in the area.

We expect the individual farmer's adoption decision to be affected by the adoption decision of other farmers in his network because farmers anticipate that they will share information with others. This knowledge may relate to the best techniques with which to grow the crop or to the profitability of the new seeds. Intuitively we expect farmers to be more likely to adopt when they know many other adopters since, due to information sharing, the marginal benefit of adopting the new seeds is likely to be increasing in the number of known adopters. Dynamic considerations, however, suggest that knowing many adopters might actually decrease the individual probability of adoption. In fact, farmers might strategically delay adoption to take advantage of the information accumulated by others in the first period. Economic theory, there-

¹See, for instance, Bertrand *et al* (2000), Besley and Coate (1992), Borjas (1992 and 1995), Case and Katz (1991), Gleaser *et al* (1996), Moffit (1983) and Woittiez and Kapteyn (1998).

²See Benabou (1996), Durlauf (1996), Lucas (1988) and Romer (1986).

³Not surprisingly, understanding the determinants of technology adoption and of the diffusion process has long been a central theme in the development literature. Early studies focus on individual and plot characteristics, see Feder *et al.* (1985) for a detailed survey.

fore, does not provide clear predictions about the relationship between individual adoption decisions and network behaviour. In particular there is no guarantee that such relationship, if it exists, is linear or even monotonous.

Our results suggest that group behaviour is an important determinant of individual adoption choices. We find that the relationship between the probability of adoption and the number of adopters known is shaped as an inverse U: the *marginal* benefit of knowing one more adopter is positive when the farmer knows few other adopters, while it becomes negative when the farmer knows many other adopters. We also find evidence that information sharing is an important channel through which group behaviour affects individual adoption choices: network effects are much stronger for farmers who report discussing agricultural practices with their network partners.

The paper makes two main contributions to the literature. First, it introduces a new data set which contains information that is needed to ameliorate the identification issues that commonly arise in this context. Second, it departs from the standard assumption of linearity⁴ and presents evidence on the precise shape of the relationship between individual adoption decisions and group behaviour, which, we argue, is consistent with a trade-off between static and dynamic incentives. Our results suggest that existing estimates are likely to be downward biased both because (i) non-linearities are not taken into account and (ii) the social network is not precisely measured.

The identification of the impact of social behaviour on individual choices is plagued by many problems that can be ascribed to two main sources: correlated nonobservables and simultaneity (see Manski (1993), Moffitt (2000), Brock and Durlauf (2000)). The problems deriving from correlated nonobservables are particularly serious when the relevant social network is not properly identified. Most studies in the literature have no direct information on the boundaries of individuals' networks and are thus forced to assume that networks correspond to some observed measure of geographical or cultural proximity.⁵ For instance, most studies on social learning in agriculture proxy networks by village averages, that is, they assume that individual behaviour is a function of the mean behaviour of other individuals in the same village. In this specification, networks are no longer individual specific, but village specific. There-

⁴Behrman et al.(2001) find evidence against linearity when estimating network effects on family planning decisions in Kenya. In contrast to this paper, however, they impose a specific structure *a priori*, i.e. they only test whether the marginal effect of knowing one adopter is larger than the effect of knowing two or more. Using employment data at the Census tract level for Los Angeles County, Conley and Topa (2002) find that informational spillovers exhibit strong decreasing returns. They assume that the information (I) received by each agent is a concave power function of the total number of employed individuals in his neighborhood (x), i.e. $I(x) = x^\xi$. Structural estimation yields $\xi = 0.38$.

⁵An important exception is Conley and Udry (2000) who have collected extremely precise data on information neighborhoods. In the labour/public literature only Woittiez and Kapteyn (1998) have direct information on the composition of the relevant social group. Munshi and Myaux (2000) and Bertrand et al. (2000) use carefully defined networks based on religion and language, respectively.

fore, social effects cannot be separated from unobservable characteristics at the village level, e.g. soil quality, that may lead to correlated behaviour within villages. Furthermore, this strategy requires all information spillovers to have the same strength. The available evidence, however, points to the contrary: in their study of information spillovers in job search in Los Angeles, Conley and Topa (2002), for instance, find that spillovers are much stronger within ethnic community than across them.

Our data allows us to improve over many of the current studies. In fact, we can identify the extension and composition of the social network *a priori*. Importantly, this allows us to control for village fixed effects, which capture any characteristic at the village level – such as soil quality, weather, or the links to traders and markets – that may drive adoption decisions.

Estimating adoption choices using cross-sectional data is complicated by the fact that there is a dynamic structure to the adoption decision (see Besley and Case (1993), Rogers (1995)). In other words, gains from adoption might depend upon the history of the new technology use up to time t (or the anticipated benefits after time t), for instance because knowledge evolves through time or because credit availability might vary. Our data is immune to this sort of problems: since we have data for the first year the innovation was introduced, there is no relevant omitted history.

Nevertheless, dynamic considerations still arise because today's decisions affect the future. In particular, a large literature both within the context of farmers' technology adoption in developing countries⁶ and of firms' and consumers' behaviour⁷ has suggested that information externalities can generate strategic delays. In our context this implies that since farmers are uncertain about the technology and since other farmers' experiences with the new crop at date 0 add to the public knowledge at date 1, each farmer has an incentive to delay adoption and then free-ride on the information accumulated by others in the previous period.

We argue that the inverse-U relationship is consistent with the trade-off between static (positive) and dynamic (negative) network effects, where the first prevail at low numbers of known adopters. In other words, if a farmer knows few adopters, the more adopters he knows the more likely he is to adopt because of information sharing at date 0. If he knows many adopters, however, the dynamic incentive to delay prevails so that knowing one more actually decreases the probability of adoption. In contrast to previous work, we are able to investigate this matter because we have precise information on networks at the individual farmer's level.

Like all studies of social interactions our analysis suffers from simultaneity bias or the "reflection problem" as discussed in Manski (1993). The issue arises because the behaviour of each farmer affects the behaviour of the other farmers in his social group. Thus the group's behaviour affects individual behaviour, which, in turn, affects the group's behaviour. We show that the severity of this problem depends on the size

⁶See Bardhan and Udry (1999), Besley and Case (1998), Foster and Rosenzweig (1995).

⁷See e.g. Caplin and Leahy (1998), Chamley and Gale (1994), Kapur (1995), McFadden and Train (1996).

of the network. We derive the bias of the OLS estimator and show that when the network is not too small⁸ endogeneity actually biases the OLS estimate of the network coefficient *downwards*.

Finally, our data allows us to control for endogenous group formation, individual heterogeneity and observable village characteristics. We show that the results are robust to such controls.

The rest of the paper is organised as follows. Section 2 describes the data set and presents the empirical methodology. Section 3 presents the results and addresses a number of econometric concerns. Section 4 presents some robustness checks. Section 5 discusses the related literature and a theoretical framework to explain the results. Section 6 concludes. All proofs are in the appendix. The data appendix gives full variable definitions.

2 Data and Methodology

2.1 Background to the Study Area

Our study area, shown on the map in the Appendix, is in the Zambezia Province, Northern Mozambique. Mozambique is amongst the poorest countries of the world, ranking 168th out of 174 countries in the Human Development Index for 2000. The country is still recovering from the turmoil of a prolonged civil war that ended in 1992. The vast majority of the population earn less than a dollar a day (per capita income in 2000 was US\$220) and around one third of national income is derived from the agricultural sector.⁹

Our data were collected in December 2000, together with the NGO *Movimondo*. The distribution of sunflower seeds¹⁰, which is the focus of the present study, is part of a larger project aimed at promoting sustainable food security by 2002.¹¹ Importantly, sunflower distribution was not conditional on participation in other *Movimondo* projects, and adoption was not a necessary condition to be able to participate to other projects.

The decision to introduce sunflower was taken for the following three reasons. First, the techniques required to cultivate sunflower are not too dissimilar from those

⁸For instance, if the social network is symmetric and contains twenty people, we show that the OLS estimate is biased downwards as long as the true coefficient is greater than 0.05.

⁹In 2000, life expectancy was 44 years, adult literacy 42%, and the infant mortality rate was 131. The figures for Northern Mozambique are worse. In addition, Mozambique was hit by severe flooding in 2000 and 2001. The most affected areas were in central Mozambique around the Zambezia river and hence do not distort our data.

¹⁰Sunflower seeds are not widely harvested in Mozambique. In 2000, only 22,000 hectares were devoted to sunflower production, less than 0.1% of the total land available.

¹¹At the time of the survey, *Movimondo* were also engaged in another project at the village level - improvements in storage technologies to reduce post-harvest losses. *Movimondo* has also been responsible for the rehabilitation of a major local bridge.

required for some of the most important existing cultures (maize, groundnuts).

Second, there is potential for farmers to commercialise sunflower production in the future by selling their production of seeds to local traders. Currently there are two principal barriers to this. First, there is a lack of legal traders (at present there are only ten legally operating traders in the district) although many more non-licensed traders and their intermediaries exist. The second barrier is that traders are unwilling to come to villages unless they can be assured of a minimum level of sunflower production. *Movimondo* is currently seeking to facilitate this process through the formation of village level farmer association groups, which should guarantee traders a sufficiently high level of production and also increases the farmers collective bargaining power. At the time of adoption, however, farmers did not know whether, or when, these groups would be established in their village.

Finally, sunflower seeds can be pressed to produce oil, which is nutritionally beneficial and currently lacking from the farmers' diets. The median farmer in our sample reports to consume oil only four days in a month. In addition, anecdotal evidence suggests that the consumption of oil is likely to raise an individual's social status within the village as oil is regarded as a luxury good in rural Mozambique.

The villages are initially contacted by *Movimondo* extension workers, who organise a village meeting and demonstration plots to illustrate the potential benefits of adoption. If a sufficiently large number of farmers in a village express an interest in the project, then *Movimondo* establishes a link with a contact farmer in each village who, in turn, supplies each farmer with his desired level of seeds. Importantly, these seeds are supplied for free. This ensures that credit constraints are not a hindrance to adoption, and it prevents the formation of a local market for sunflower seeds.

2.2 Data

The sample was chosen in a two-step cluster process - first choosing the villages, then households. No census is available for the study area hence it is not possible to claim that the sample is representative of rural Mozambique. Instead, it was decided to collect the sample to have a representative sample of villages in terms of village wealth and soil quality.¹² Nine villages were selected in total, covering 204 households.¹³ Within these villages, a random sample of adopters was chosen (from *Movimondo* records), and we then aimed to randomly survey an equal number of non-adopters. *Movimondo* extension workers who had been working in each village

¹²None of the villages have adequate infrastructural links to the provincial capital, Nampula. There is no tarmac road in the study area, although part of the tertiary road system has recently been rehabilitated, bridge infrastructures remain weak. During the rainy season a large proportion of the study area remains unreachable by road since several bridges remain under water. In our data, the average travel time (on foot) to the nearest permanent market was 76 minutes, and to the nearest road was 43 minutes. See table 1.

¹³In addition two villages where *Movimondo* will begin supplying sunflower in 2001 were also sampled. These are not used in the regression analysis.

then administered the questionnaire. Further details of the study area, questionnaire, and the sampling methodology can be found in Gallego-Lizon and Rasul (2001).

Reasons for (not) Adopting.

The first part of table 1A shows that most farmers decided to adopt the new crop in order to consume the oil. As explained above, commercialising the seeds is currently very difficult and indeed very few farmers reported having done so. Interestingly, many of the non-adopters indicate that they chose not to adopt because they were not familiar with the techniques required for the new crop. The data also suggests that heterogeneous crop preferences are not an important determinant of the adoption decision: only 3% of farmers report not adopting because they did not want to give up their existing crops.

To have some direct evidence on the importance of peer effects, we asked the farmers who lived in villages where the seeds had not been yet introduced whether they would be more willing to adopt if others were adopting at the same time. The answers, reported in the table B, clearly suggest that peer effects are important. Overall 92% of the thirty-seven farmers we interviewed, said they would be more willing to adopt if either a friend, a family member or a fellow church member would adopt at the same time. Interestingly, peer effects do not cross the village borders: only one farmer reported being more willing to adopt had his friends in other villages adopted as well. This may be because such information is more costly to verify.

Demographic Data

Table 1 gives a complete list of descriptives by adoption status. We see that the groups do not differ significantly by observable characteristics such literacy, age and income.¹⁴ Female headed households are more likely to adopt. We have data on the number of months of the year for which the household is food secure, i.e. the household has food stocks available in those months. Table 1 shows that adopters and non-adopters do not differ in their exposure to food insecurity.

We also see that the groups of adopters and non-adopters are almost identically distributed across poverty groups (defined through asset ownership - see the data appendix). This is as we would expect given that the seeds are distributed for free and hence we would not expect credit constraints to hinder adoption.

Measures of Unobservable Individual Heterogeneity

We have a number of measures which may proxy for unobservable heterogeneity amongst farmers. Examples of such unobservables include ability, entrepreneurship, and credit constraints. We have information on whether the individual has ever participated in any other NGO projects in the past. The majority of villages in the village have had other NGO projects at some time. Also, we asked individuals if they were aware of local agricultural prices, and if so, from which source did they obtain

¹⁴Throughout the paper we assume households follow a unitary decision making process. To support this assumption, note that 48% of spouses are reported as having their own plots. In addition for those households in which the spouse do have separate plots, none of these spouses adopt when their partner does. For non adopters the figure is 5%, which is not significantly different.

this information (themselves, family members, neighbours/friends, or members of their church).¹⁵ We use this to construct a dummy equal to one if they obtained this price information themselves. Finally, we asked the respondents “who do you speak to about agricultural decisions?”. Again, we construct a dummy equal to one if they speak to at least one other source. Table 1 shows that adopters are significantly more likely to have participated in NGO projects in the past, have obtained agricultural price information themselves, and speak to others about agriculture in general.

Social Networks Data

We asked individuals “how many people do you know that are planting the new cultures?” in the following categories - family, fellow church members, neighbours, friends in other communities, and others. We use the sum of these as a measure of the aggregate size of the individual’s social network. The descriptives given in table 1 show that adopters know significantly more other adopters than non-adopters.¹⁶ Breaking this down by the three most important groups (family, neighbours and church members), we find, as expected, that adopters are significantly more likely to know adopters in each of these categories. The descriptives suggest that there are largely no differences in the types of people in the social networks employed by male and female household heads. There is some weak evidence suggesting the females are more likely to know family members, and knowing more neighbours adopting than males.

On face value, these descriptives suggest that unobservable rather than observable characteristics drive adoption decisions, and that adopters have greater social networks than non-adopters. These networks may differ also by gender of household head. The regression analysis later in this section addresses whether these partial correlations are robust to controlling for the other covariates.

Village Characteristics

In addition to the household questionnaire, we also conducted a separate survey with three village elders in each village, in order to obtain village level information.¹⁷ The third part of Table 1 reports whether any other NGOs also operate in the village,

¹⁵Throughout the analysis, whenever we refer to neighbours we are referring to neighbours or friends within the village.

¹⁶One concern is that 18% of the sample report not being numerate. As a check, we also asked individuals the same question at a different part of the questionnaire but where the enumerator offered choices of none, 1-5, 5-10, 11-20, 21-30, more than 30, in each of the categories. We then constructed an alternative measure of the aggregate social network based on the midpoints of these bins. We found the correlation between the two measures to be 0.95. The correlation is not significantly different between those who can and cannot count.

¹⁷For want of a better expression we refer to these individuals as village elders, although they are not necessarily above the average age of respondents in the village. The three elders in each village are a traditional leader, the contact farmer (these are the farmers who act as intermediaries between *Movimondo* and the other farmers in the village), and a political leader. The three leaders gave largely consistent information. Wherever possible, we use the answers of the traditional leader. Where there is missing data, we use the responses of contact farmers. The contact farmers are not in our sample for the individual level analysis.

what infrastructure exists in each village and the distance to the main road. We asked village elders a number of questions relating to the social insurance mechanisms available to individuals, including “what coping strategies do people use if they have a shortage of food?”. This question allows us to distinguish across villages where individuals are largely self reliant, reliant on others within their own village only, or reliant on both those in their own village and from other villages. Table 1 shows that villages do differ in the levels of infrastructure present, and in the sources of social insurance. Some villages such as Inlepa appear to have good infrastructure and availability of outside insurance, whereas others such as Morrua lack some of the most basic forms of village infrastructure, and no forms of insurance are available to members in times of food shortages. This creates enough variation in the data across villages for us to separately identify the effects of village infrastructure and the sources of social insurance, on adoption decisions. We turn to this in the next section.

2.3 Methodology

Following the large body of theoretical and empirical literature on peer effects and social networks, considerable attention has been devoted to the econometric problems that typically arise in this context. The first part of this section discusses identification issues in a general setting while the second part relates them to our specific context and presents the methodology employed in this study.

The literature on peer effects and social interactions aims to identify the effect of group behaviour on the behaviour of the individuals that belong to the group. Let a_{iv} denote the behaviour of individual i living in area v , and let $n(i)$ denote the social group of individual i . If one believes that social interactions affect individual decisions, a_{iv} should be estimated as a function of the personal characteristics of individual i (X_i), and the characteristics of the relevant social group i belongs to (both their actual behaviour, $a_{n(i),v}$, as well as their characteristics, $X_{n(i),v}$). Assuming linearity we have;

$$a_{iv} = \beta X_i + \gamma a_{n(i),v} + \delta X_{n(i),v} + u_{iv} \quad (1)$$

where u_{iv} is an error term assumed to be uncorrelated with any of the other covariates as the model is correctly specified.

Using the terminology of Manski (1993), γ measures the “endogenous” social effects, i.e. the impact of group behaviour on the choice of individual i while δ measures the “exogenous” social effects, that is the direct effect of the characteristics of the group on the decision of i .

An example of an endogenous social effect in our study is if individual i is more likely to adopt if more members of his network adopt because once they adopt, i believes his social group will be able together to solve technical problems that arise, or advise one another on the optimal input choice to use and so forth. Hence the

endogenous social effect operates through any channel in which the fact that members of the network have *also* adopted, will raise the marginal benefits of i himself adopting. This is one reason why behaviour should therefore be correlated across the social group.

An example of an exogenous social effect in our study is where i 's social group comprises of individuals that have some characteristic, say the ability to read, or their connections to local traders, which raises the marginal benefit to i adopting irrespective of whether the members of i 's social group have themselves adopted. This provides a potential reason why the actual behaviour of a social group may be correlated, but it is the characteristics of the group as opposed to their actual behaviour which is the causal mechanism for this correlation.

Although difficulties in the identification of "social" effects are many and diverse and depend on the particular context of the study, they all derive from two root sources - simultaneity and correlated unobservables.

Simultaneity arises because social interactions are symmetric in the sense that i 's behaviour affects the behaviour of the other members of the group (i.e. $a_{n(i)v} = g(a_{iv})$). This is known as the reflection problem - group behaviour affects i 's behaviour, which, in turn, affects the group's behaviour. Due to simultaneity, the parameters (β, γ, δ) in (1) are efficiently estimated in a system of equations, with corresponding equations for all $j \in n(i)$. It can then be shown that γ and δ - that is the "endogenous" and the "exogenous" social effects - are separately identified only if the errors of the structural form (that is u_{iv} and u_{jv} for all $j \neq i, j \in n(i)$) are uncorrelated. To see this, consider the simplest case in which $i = 1$ and $n(i) = j = 2$, i.e. each farmer's social group consists of only one other member. The structural form of the model is then;

$$\begin{aligned} a_1 &= \beta x_1 + \gamma a_2 + \delta x_2 + u_1 \\ a_2 &= \beta x_2 + \gamma a_1 + \delta x_1 + u_2 \end{aligned} \quad (2)$$

where $\text{cov}(u_1, u_2) = 0$ as the model is assumed to be correctly specified. The reduced form is given by;

$$\begin{aligned} a_1 &= b x_1 + c x_2 + v_1 \\ a_2 &= b x_2 + c x_1 + v_2 \end{aligned} \quad \text{where } b = \frac{\gamma\delta + \beta}{(1 - \gamma^2)}, c = \frac{\gamma\beta + \delta}{(1 - \gamma^2)}, v_1 = \frac{u_1 + \gamma u_2}{(1 - \gamma^2)}, v_2 = \frac{u_2 + \gamma u_1}{(1 - \gamma^2)} \quad (3)$$

Thus γ can be recovered from the covariance terms, and then the estimates of b and c used to recover β and δ .

The assumption of uncorrelated errors, however, is appropriate if and only if the decisions of i and $j \in n(i)$ are correlated exclusively because of endogenous or exogenous social interaction. In practice, i 's behaviour can be correlated with j 's for

a number of other reasons – which leads to the second of the identification problems listed above, namely correlated unobservables.

Correlated unobservables can be of two types, depending on whether they are contextual –i.e. specific to the dimension along which groups are defined–or individual specific. The first type are variables that affect every member’s adoption decision and are specific to the environment. For instance, if $n(i)$ is defined as (or proxied by) the population in a village, village unobservables such as soil quality and credit markets make i ’s decision correlated with the group’s even if social interactions are altogether absent.¹⁸

Variables of the second type are specific to the individual and determine both the individual outcome and the decision to join a given group. It follows that, because of individual unobservables, group membership itself is endogenous. So, for instance if more entrepreneurial farmers are more likely to adopt a new technology and also more likely to be in the same social group as other entrepreneurial farmers, the individual and the group outcome would be correlated–again even in the absence of social effects.

To summarise, to be able to identify endogenous and exogenous effects in the Manski (1993) framework, the econometrician is required to have the following data - (i) a complete set of individual characteristics that determine adoption. This includes variables that we typically think of as being unobservable such as farmer ability, entrepreneurship, “openness” towards new ideas, or credit constraints; (ii) the behaviour of the true social group of i , $n(i)$ which of course assumes that $n(i)$ has been correctly identified; (iii) the relevant set of characteristics of the relevant social group, $X_{n(i),v}$. Needless to say, these data requirements are very stringent, and not typically found in data sets unless they have been specifically collected to study social effects. In the absence of any one of these sets of controls, the identification of social effects becomes confounded with the correlated unobservables problem discussed above.

Most of the existing literature is restricted by not having such rich data available, hence inevitably a number of simplifications have to be made. First, individuals typically do not report their social groups. Rather, social groups are proxied by geographical or cultural proximity.¹⁹ Hence it is assumed that the behaviour of individual i in village v , say, is a function of the mean behaviour of individuals in the village, \bar{a}_v . The important point to note is that is no longer an individual specific determinant, but village specific. This makes it impossible to separate out social effects from unobservable characteristics at the village level that may lead to correlated behaviour within villages.

Second, as a consequence of having to proxy the true social group of i with his

¹⁸Examples of environment-specific omitted variables in the literature on peer effects on educational achievement include teacher quality and school resources when the social group is defined as peer students in the same class.

¹⁹For instance, in the literature on neighborhood effects in cities individual outcomes are estimated as a function of outcomes of other individuals in the neighborhood. That is, it is implicitly assumed that the relevant social group consists of neighbours only.

village, it is also not possible to control for the common characteristics of the group which may drive adoption decisions, namely, exogenous social effects cannot be identified separately from endogenous social effects. In short, much of the literature has estimated the following type of specification for individual i in village v ;

$$a_{iv} = \beta X_i + \gamma \bar{a}_v + u_{iv} \quad (4)$$

From our discussion it should now be clear that in this specification, a positive γ coefficient cannot simply be interpreted as providing evidence of any social effects - it could be simply due to common unobservables. Even in the absence of such effects, we cannot disentangle endogenous from exogenous social effects.²⁰

Our data helps us to make two improvements over many of the current studies. First, we specifically asked each individual about their social network. Thus we can identify the extension and composition of the social network *a priori*. We do not need to make assumptions about the level at which the group is defined. In other words, $n(i)$ is identified in our data. This allows us to also control for village fixed effects, so we control for any characteristic at the village level that may drive adoption decisions, such as soil quality, weather, the availability of credit, links to traders and so forth.

Whilst we do have some variables that may be proxies for sources of heterogeneity across individuals that both drives individual adoption decisions, endogenous group formation is still an issue in this context and will be addressed in the section on robustness checks.

To summarise, we estimate the following equation;

$$a_{iv} = \beta X_i + \gamma f(a_{n(i)v}) + \delta_v + u_{iv} \quad (5)$$

where $a_{iv} \in \{0, 1\}$ represents the individual adoption decision, X_i is a vector of individual characteristics, $n(i)$ is the number of people in the same social group as i and δ_v are village fixed effects. We assume that i 's decision depends on the adopters in his network and employ two alternative definitions for $a_{n(i)}$, namely the total number of adopters in i 's network ($\sum_{j \in n(i)} a_{jv}$) and a dummy variable that equals 1 when there is at least one other adopter in the network. While the first is our preferred specification, we also present results for the second as a robustness check in case the continuous measure of the number of adopters known is subject to measurement error.

Rather than assuming a particular functional form for $f(a_{n(i),v})$, we use non-parametric methods to guide our choice of f . We then use a linear approximation for f and estimate (2) by linear probability. Although this model might yield predicted values outside of the feasible $[0, 1]$ range, it is preferable to non-linear models mainly

²⁰The fact that we cannot disentangle endogenous from exogenous social effects would not be aided even if we had information on $n(i)$ and $X_{n(i),v}$. To see this, note that estimating (2) in reduced form, then given that $\beta \neq 0$, if $c \neq 0$ then this can arise because endogenous social effects exist ($\gamma \neq 0$) and/or because exogenous social effects exist ($\delta \neq 0$). Hence, we can only identify if social effects exist *per se*, but not the causal mechanism through which they may effect group behaviour.

because we can then sign *a priori* a number of important sources of bias such as omitted variables. Moreover, predicted values generally belong to the $[0, 1]$ interval as long as the mean of the dependent variable is not too close to 0 or 1, which is true in our case. The mean of our dependent variable, whether a farmer adopts or not, is 0.525.²¹

What our data lacks however is the characteristics of i 's social network, $X_{n(i),v}$. This, in turn, precludes the use of system analysis to separate direct "exogenous" social effects from indirect or "endogenous" ones. Even if system estimation is not feasible, we argue that it is possible to identify social effects, whether direct or indirect, from pure reflection. That is, we can show that abstracting away from issues of non-linearity and assuming all individuals have the same social network of size n , then if the true data generating process is given by;

$$a_i = \beta X_i + \gamma \left(\sum_{j \neq i}^n a_j \right) + \delta \left(\sum_{j \neq i}^n X_j \right) + u_i; \quad i = 1, \dots, n$$

but the econometrician, due to data constraints, estimates the following model;

$$a_i = \beta X_i + \gamma \left(\sum_{j \neq i}^n a_j \right) + u_i; \quad i = 1, \dots, n$$

It is implicitly assumed that the true parameters of the DGP, (β, γ, δ) are all non-negative. If so, then the following result applies to the OLS estimate of the endogenous social effects in the misspecified model, $\hat{\gamma}$.

Proposition 1 *In the absence of correlated effects ($cov(u_i, u_j) = 0$ for all $i \neq j$):*

- (i) *if there are no true social effects ($\gamma = \delta = 0$), $\hat{\gamma}$ is consistent*
- (ii) *with social effects, $\hat{\gamma}$ is inconsistent. A necessary and sufficient condition for $\hat{\gamma}$ to be downwards biased is if;*

$$\gamma > \frac{1}{n-1}$$

where n is the size of the social network. With correlated effects ($cov(u_i, u_j) = \rho$ for all $i \neq j$):

- (i) *if there are no true social effects ($\gamma = \delta = 0$), $\hat{\gamma}$ is inconsistent with $plim \hat{\gamma} > 0$*
- (ii) *with social effects, $\hat{\gamma}$ is inconsistent. A sufficient condition for $\hat{\gamma}$ to be downwards biased is*

$$\gamma > \frac{2}{n-2}$$

²¹This is obviously a direct consequence of our sampling strategy. The actual adoption rate in the study area is considerably lower.

In the case of uncorrelated errors, the condition required for $\text{plim } \hat{\gamma} < \gamma$ is not very stringent. For instance, if the true social network consists of 20 individuals, the true γ ought to be greater than 0.05 for our estimate to underestimate it. In the second case, the conditions required for downwards bias in our estimated coefficient still hold over a wide range of (n, γ) pairs that we may *a priori* consider reasonable.²²

3 Empirical Analysis

3.1 Main Findings

Tables 2 and 3 report our main findings. We first analyse the effect of individual characteristics and village mean adoption rates as in specification (2), the “classic” variable commonly employed in the literature to proxy for social networks. We then employ our measures of individual networks and analyse whether the data supports the idea that knowing more adopters increases the probability of adoption. In the same context, we explore whether the relationship is linear. Finally we attempt to assess the extent to which network effects operate through information sharing.

3.1.1 Baseline Specifications

Column 1, Table 2 estimates individual adoption decisions as a function of individual characteristics, which can be divided into demographics, wealth/risk measures and proxies for attitudes/entrepreneurship. To select the relevant individual characteristics we follow the large literature on the diffusion of innovations (Rogers 1995), which suggests that education, wealth, exposure to risk and exposure to media are significant determinants of early adoption. Food security, a measure of risk exposure, does matter for adoption: farmers who are more exposed to risk – less months of food security – are less likely to adopt. Wealth, measured by relative asset holdings, is not a significant determinant of adoption, which is consistent with the fact that, as discussed above, credit constraints are not an issue in this context.²³ That wealth constraints do not matter is further confirmed by the fact that female headed households, among the poorest in our sample, are actually more likely to adopt. Female

²²In the discussion above we have focused exclusively on symmetric networks, i.e. networks in which everybody knows everybody else. Existing sociological and anthropological evidence however suggests that networks can take many other shapes. The nature of the reflection problem and the bias of the OLS estimator depend on the exact structure of the network. (see Bandiera and Rasul (2002)).

²³The results are robust to the exclusion of the food security variable and to different poverty definitions. Food security does not significantly differ across poverty groups, hence it is unlikely to be acting as a proxy for wealth. We experimented with a number of baseline specifications. Throughout we found the respondent’s education, religion, roof or wall materials all to be insignificant. In addition we do not control for tribe as 99% of our sample belong to the Lomue tribe. This is typical of the region. Finally to capture exposure to media we checked whether having a radio affected the probability of adoption. We found no significant impact.

headed households, literate and, to a certain extent, older farmers are more likely to adopt. Past participation in other NGOs projects is included to proxy for individual personality traits which make some better inclined towards these kind of ventures.²⁴ We find, as expected, that past participation in NGO projects has a strong positive effect on the probability of adoption. Direct knowledge of agricultural prices separates farmers who have invested resources into finding information for prices from farmers who either do not know produce prices or know them through indirect sources. Direct knowledge of prices is a proxy for farmers' entrepreneurship and ability/opportunity to commercialise crops. Its effect is positive and significant. Finally, farmers who discuss agricultural issues with others are also more likely to adopt. It is important to note that all individual variables date before the introduction of sunflower seeds and are therefore exogenous to the adoption decision.²⁵

Column (2) includes the "classic" measure of network effects, that is, mean adoption rate at the village level. This specification thus replicates what much of the earlier literature on peer effects has estimated, in the form of specification (2) in the earlier section. The results suggest that farmers who live in villages where the adoption rate is higher are more likely to adopt, the coefficient is positive and marginally significant. Column (3) adds our measure of social networks, that is the number of adopters known by each farmer. Individual networks have a positive and significant effect, even controlling for mean adoption rate at the village level suggesting, as suspected, that individual networks matter over and above any village level determinants.²⁶

There are two ways to interpret this result. First, it may suggest that some information within the village is public, so that village adoption rates measure the amount of this public information available to each farmer in order for him to solve any problems that may arise after adopting. In addition there is also information which is private to i within i 's social network, and this is picked up by the coefficient on the number of adopters known to i .

²⁴That is, non *Movimondo* projects previously operating in the village.

²⁵Land size is not measured in the data since farmers themselves do not know it. Omitting land size, however, should not invalidate the results because land is abundant in these villages. Moreover, landholdings should be correlated with wealth, which is not a significant determinant of adoption.

We are not able to control for plot level soil quality either. This concern is partly ameliorated by two factors. First, land markets are missing. Land is mostly inherited, or obtained through the permission of village elders. Hence it is unlikely that more able farmers would necessarily be able to obtain better quality soil. Second, plots tend to be spread out in different parts of the village. If soil nutrients are randomly distributed in the village, this would reduce the variation in average plot quality across farmers within the same village.

²⁶A potential source of concern is that, since we do not have a complete map of village networks, we are unable to say whether farmers belong to the same network within one village. To the extent that they do, the standard errors are biased downwards because observations from the same group receive the same weight as observations from different groups. Ideally, we should cluster at the network level. Instead, due to data limitations, we re-estimated the model correcting for clustering at the village level—which is actually more stringent. The results, not reported for reasons of space, are unchanged.

Of course, the second way to interpret the result is that there is some village level characteristic that village adoption rates proxy for (say better quality soil) that has nothing to do with social effects. The results are more in line with the latter interpretation: indeed, if the village average were proxying for network effects, its coefficient should fall once individual networks are included. A comparison of columns (2) and (3), however, reveals that the coefficient on village average does not change as a result of the inclusion of the individual network variable. The results thus suggest that only the information which i can obtain through his network can be useful to him. We go on to examine this in more detail in a later section.

3.1.2 Non-linearities.

The next column shows that the relationship between the probability of adoption and the number of adopters known is non-linear. Here we drop the village mean effect and introduce village fixed effects, to control for village unobservables which could affect the adoption decision.²⁷

In table 2, column (4) we introduce the square of the number of adopters known, which is negative and significant, suggesting decreasing returns.²⁸ Importantly, the results suggest that assuming linearity leads to seriously underestimate the network effect. The coefficient on the number of known adopters is five times larger when the square term is introduced (see columns (3) and (4)).

Parametric regressions, however, provide only limited information on the shape of the relationship between the variables of interest. To explore the matter further we employ non-parametric methods. We first regress;

$$a_{iv} = b + cX_i + d \sum_{j \in n(i)} a_{jv} + \delta_v + e_{iv}$$

and then use the predicted values of the coefficients to compute $a_{iv} - \hat{a}_{iv}$, where $\hat{a}_{iv} = \hat{b} + \hat{c}X_i + \hat{\delta}_v$, and then run a kernel regression of $a_{iv} - \hat{a}_{iv}$ on the number of adopters known. Note that by using $a_{iv} - \hat{a}_{iv}$ instead of a_{iv} , we control for the variation in a_{iv} explained by individual and village characteristics. The kernel regression with 95% bootstrapped confidence intervals, is shown in figure 1. This again suggests that the relationship between the probability of adoption and the number of adopters known is clearly non-linear: a_{iv} is first increasing and then decreasing in the number of adopters known. In other words, the probability of adoption is low both when the farmer knows very few or many other farmers who have adopted.

The result is not driven by outliers, about 20% of farmers lie to the right of the maximum point in figure 1, that is they know more than thirteen adopters.

²⁷Village effects, not reported for reasons of space, are jointly significant.

²⁸The fact that the relationship is non linear also allows us to dismiss the argument that knowing many adopters might be proxying for the total size of the network, which could be correlated with other characteristics, e.g. openness, that make adoption more likely.

However, since the analysis relies on a small sample, especially for farmers who know more than twenty adopters, non-parametric estimates cannot be very precise. As further check we estimated a piecewise linear function where we allow the coefficient to be different as suggested by the non-parametric estimates. The parametric results, column (5) Table 2, are in line with the previous estimates: knowing between 10 and 16 adopters has a significantly stronger impact on the probability of adoption than knowing between 1 and 9 or more than 16.

One possible explanation for the result is that people who know many adopters, i.e. more than thirteen in our case, are systematically different along a dimension that reduces the probability of adoption. For instance, these could be the most entrepreneurial farmers who are already engaged in some other profitable enterprise and therefore not interested in cultivating sunflower. Descriptive statistics do not show significant differences in the relevant observable dimensions such as wealth, past participation to other NGO activities, education and most other variables. People who know many adopters, however, are more likely to have cashew trees and to grow (and consume) more vegetables and fruit. To check the robustness of the results we added a cashew and fruit dummy in the regressions, which, however, were never significant and did not affect the other coefficients.²⁹

3.1.3 Does Information Sharing Matter?

We now attempt to assess whether network effects operate through information sharing or through other channels. To do so, we test whether network effects have a different impact on the probability of adoption for farmers who discuss agricultural practices with others compared to farmers who do not. Intuitively, if the impact is the same, sharing information about agriculture cannot be the mechanism through which network effects affect the probability of adoption. Columns (2) and (4) of Table 3 include an interaction term between our network variables and a dummy which identifies whether a farmer talks to others about agriculture. In both cases the interaction term is positive and significant: networks effects are stronger for farmers who share information about agriculture. For instance, using the results in Column (2) we see that the marginal effect of an additional member of the farmer’s social network adopting (in the sense that i knows that this farmer is adopting), is to raise i ’s adoption probability by 4.6%. In addition, if i talks to farmers about agriculture, this marginal effect rises by an additional 3%. Hence network effects are stronger for more “talkative” farmers. Relatedly, this shows that although mimicry might be playing a role, we can safely rule it out as the main determinant of our estimated social effects. Indeed if mimicry were the only factor, the coefficient of the network variable should be the same for farmers who share information and for farmers who

²⁹Note that since both fruit and cashew are trees endogeneity is not an issue. That is, both fruit and cashew were planted before the introduction of sunflower. We cannot claim the same for vegetables and hence do not include a vegetable dummy in the regressions.

do not.

Columns (3) and (4) are identical to (1) and (2) except now we use the dummy variable of whether the farmer knows anyone at all adopting to capture his social network. Again the interaction effect is positive and significant. The fact that the dummy variable of knowing at least one farmer is not significant in this specification is not altogether surprising given the earlier evidence that such network effects are non linear.

To conclude then, overall our results suggest that information sharing is an important determinant of the adoption decision. In short, social effects do matter. However with our data, we cannot separate out the endogenous social effects from the exogenous social effects. Also, there are a host of econometric concerns and robustness checks to which we now turn.

3.2 Main Econometric Concerns

Following the analysis in section 2.3 we can classify econometric concerns according to whether they stem from correlated unobservables or from simultaneity.

3.2.1 Unobservable Heterogeneity

Regarding correlated unobservables, we have already argued that contextual (e.g. area specific) nonobservables are not an issue in this context because our network variables are based on the actual network rather than on groups that are geographically or culturally close to the individual farmer. We have controlled for village level unobservables using village fixed effects throughout. Individual omitted variables, however, are still a cause for concern.

These are of two types. First, there are individual attributes which affect both the adoption decision and the number of adopters known. Second, group membership might be endogenous in the sense that farmers might have a personality trait, for instance “smartness”, or a common preference to consume oil, which increases the probability of adoption and at the same time makes them want to group with farmers with the same trait.

Nonobservables of the first type include entrepreneurship, openness towards new ideas, extroversion and the like. For instance, more entrepreneurial farmers are likely to know more people for their business and at the same time more likely to adopt. To address this issue we include individual “attitudinal” variables in all regressions. These variables measure past participation in other NGO activities (a proxy for openness to new ideas and also good attitude towards NGOs in general), whether the farmer has direct knowledge of agricultural prices (a proxy for entrepreneurship and commercialisation) and whether the farmer talks to others about agriculture—a proxy for extroversion and friendliness. To assess whether these variables effectively capture some of the unobservable heterogeneity that would otherwise bias the social network

coefficients upward, we compare the social network coefficients that we obtain with and without the heterogeneity controls.

The results are reported in Table 4, columns 1-3 for the continuous network variable-columns 4-6 for the network dummy. The models in columns 1 and 4 estimate the probability of adoption as a function of network effects and village fixed effects only. The models in columns 2 and 5 add exogenous individual characteristics (age, gender, education etc.). The models in columns 3 and 6 add the heterogeneity controls. Note that the coefficient on social networks does not change from 1 to 2 (respectively 4 to 5), i.e. when exogenous controls are included. This is as expected as these exogenous variables should not be correlated to the network effects. However, the coefficients on social networks in both specifications fall when the heterogeneity controls are included (by 25% for the continuous variable). This suggests that, as suspected, the network coefficients would be overestimated if heterogeneity controls were not included and that, most importantly, the controls we employ are, in part, proxying for unobserved individual characteristics.

3.2.2 Endogenous Group Membership

Nonobservables of the second type are individual specific characteristics which increase the probability of adoption and at the same time determine group membership. Since we do not have data on other group members we cannot find direct proxies for this kind of omitted variables. The issue of endogenous groups, however, arises only to the extent that individuals are allowed to choose their network. Based on this idea, we divide the number of adopters known into three groups: (1) number of adopters known in the (extended) family, (2) number of adopters known in the same church, (3) number of adopters known who are neighbours/friends. Amongst these three groups, we argue that as a farmer cannot choose his family, and to a lesser extent cannot choose his fellow church members, any self selection into a social group can only occur through your choice of known neighbours.³⁰

If endogenous matching were to be the only driving force behind our estimated network effects, we should find that the coefficients of groups (1) and (2) (family and church) are not significantly different from zero. The estimates in columns 1-3 of table 5 show that this is clearly not the case. In column (1), knowing family members, as well as neighbours, who also adopt, increases your own adoption probabilities. In column (2) where we allow for non linearities in the social effects for each group, we find diminishing returns to knowing members of your own family and neighbours, on your adoption decision. To address the issue of whether male and female headed households may base their networks amongst different groups, we interacted each

³⁰Church membership would be endogenous if religion was correlated with such unobservables and individuals changed religion to be with people with similar traits. This appears highly implausible to us. In addition, choice of church effectively implies choice of religion as there is only one type of each church in each village. Hence we do not observe sorting of types across churches within a village.

group with a gender dummy. Column (3) shows that females are less likely to know fellow church members adopting, but the social effects operating through family and neighbours are identical across genders.³¹

As a further check we propose another test. The test is based on the following intuition. Suppose there is some unobservable variable that determines group membership and individual adoption probability. For example, this may be a preference to consume oil, so that all oil lovers know one another, and similarly for oil haters. In this case, the average number of adopters known to other adopters in i 's village would proxy for the tastes for oil in the adopting group. Similarly, the average number of adopters known to non adopters would proxy for the non adopting group's (dis)taste for oil. We denote the number of other adopters known to the group (either the adopters or non-adopters) that i belongs to as being $\overline{a_{iv}}$, and then estimate;

$$a_{iv} = \beta X_i + \gamma f(a_{n(i),v}) + \theta \overline{a_{iv}} + \delta_v + u_{iv} \quad (6)$$

If farmers form social groups on the basis of their preference for oil, then their own social network and the average network of other members of their group ought to be coincident and (γ, θ) are unidentified as $a_{n(i),v} = \overline{a_{iv}}$. Taking a less extreme position where tastes for oil is one reason for group formation which also drives adoption decisions, then (γ, θ) are both individually identified³². However, we ought to find that θ picks up any social effects due to common preferences, and γ then picks up any additional effects of group membership (related to any other factor).

Column (4) in table 5 repeats our earlier preferred baseline specification for comparison. In columns (5) and (6) we additionally control for the average number of adopters in i 's group (be it other adopters if i adopts, and non adopters if i does not adopt). In column 6 we drop the village fixed effects. We find that θ is both positive and significant, suggesting that some common unobservable such as the taste for oil, does drive group formation and adoption decisions. However, the number of adopters known to i continues to have a positive and significant effect. This social effect still has diminishing returns. One way to interpret this is that the social group known to i has an effect over and above the average group known to other adopters (or non-adopters if i is himself a non-adopter). Hence individual social networks still matter in addition to any unobservable that drives group formation.

³¹In columns (1) - (3) the coefficient on knowing neighbours is never significantly greater than that on knowing family members. This again would suggest that endogenous group membership is not an issue.

³²The sample correlation between $\sum_{j \in n(i)} a_{jv}$ and $\overline{a_{iv}}$ is 0.53.

4 Robustness Checks

4.1 Regression Specification

Throughout our analysis we have employed the linear probability model to estimate our parameters of interest. Given that we have a dichotomous dependent variable, it is more efficient to use a probit estimator. Table 6 does this for our four main specifications. Again we find that the relationship between the probability of adoption and the number of known adopters is shaped as an inverse U: the square term is positive and significant and the piecewise regression shows a maximum effect between 10 and 16. In line with previous estimates, network effects are stronger for farmers who discuss agriculture with others and endogenous group formation does not seem to be a concern.

4.2 Simultaneity

To address the issue of simultaneity and, relatedly, endogenous group formation and exogenous social effects, we seek instruments for the number of adopters known. The instruments we seek ought to be correlated with the social group of farmer i , but uncorrelated to any unobservable that drives both group membership and i 's adoption decisions. If our instrumentation procedure is valid we should obtain consistent estimates of the social effect.

In our survey we asked questions to serve as such instruments. In particular we asked respondents, to list the sources, if any, they had obtained information from with regards to two important local events - first, the electrification of the main town in the sample area, Gile, and the rehabilitation of a bridge in the Gile District, conducted by *Movimondo*. We then classified the individuals as either finding out about the event by themselves, through others, or being unaware of the event. In addition we use the number of different sources of information about each event as instruments for the number of adopters known.

The idea behind the instruments is that the farmers who did not know about these non-agricultural events were more isolated from the community, hence unlikely to know many people in general, while farmers who discussed the news with friends/family were better integrated. Table 7 reports the first and second stage instrumental variables regressions. In the first stage, the instruments are reasonably good predictors of the number of adopters known. The F-test of under-identification of the instruments has a p-value of zero.

The second stage regressions show that after instrumenting, the coefficient of network effects remains positive and significant. Moreover, the coefficient is twice as large after instrumenting, suggesting that indeed endogeneity creates a downward bias in this context. The instrument set passes the Lagrange multiplier test of overidentification.

4.3 Omitted Village Characteristics

The fact that the village average adoption rate is significant even after including the measure for personal networks, combined with the fact that village fixed effects are always significant suggests that village characteristics must be important determinants of adoption. The purpose of this section is to identify some of them.

The village questionnaire contains information on a host of village characteristics, such as infrastructure—schools, health centres, wells, food markets—, distance to main road, presence of other NGOs and available insurance mechanisms against food shocks. We estimate our basic specification with these village characteristics using both linear probability and probit regressions. The results, reported in Table 8, show that farmers who live in villages that are closer to the main road are less likely to adopt. This might be due to the fact that, since most farmers report consumption as the main reason for adoption, farmers who live by the road have better access to other sources of oil (e.g. traders or markets). A quick glance at the oil consumption data does indeed suggest that, before the introduction of the new seeds, farmers in better connected villages consumed oil more often.³³ Table 8 also shows that farmers who live in villages where another NGO is active are less likely to adopt, suggesting that different NGOs can be substitutes to an extent. Village infrastructure and access to outside insurance are not significant determinants of adoption. Finally, as explained in section 2 above, *Movimondo*, runs the programme through a contact farmer in each village. The contact farmer holds the seeds and distributes them to the other farmers. It could be argued that the personal characteristics of the contact farmer matter for the probability of adoption in each village. To check for this we controlled for contact’s age, religion, participation to past NGO projects, knowledge of agricultural prices and radio ownership. Of these only the latter had a positive and significant effect on the probability of adoption. We also checked whether farmers who are in the same age group or belong to the same religion as the contact farmer were more likely to adopt. We found no evidence in favour of this effect.

4.4 Alternative Poverty Measures

Throughout our analysis we have argued that as the sunflower seeds are distributed freely to farmers, there ought to be no credit constraints to adoption. This helps us to eliminate one of the main sources of unobservable heterogeneity that may have spuriously led to us finding the existence of social effects. Throughout our regressions we have controlled for the asset poverty status of the household. Nowhere did we find this measure to be a significant determinant of adoption decisions. As a robustness check, we replace this poverty measure with various alternative poverty measures based upon animal ownership, consumption of food, log(income), and a

³³Ordering villages from the closest to the farther (from the main road), the average (median) oil consumption in days per month is: 15(12), 9.4(4), 7 (4), 6.5 (4), 8.4 (4), 2 (0), 1.7 (0).

more subjective measure based upon the questionnaire enumerator’s own assessment of the poverty in the household. The results are reported in table 11, where we find that none of these measures are significant determinants of the adoption decisions. With each of them, we continue to find evidence for non linear network effects.

4.5 Imitation Effects

One of the major difficulties in studies of group behaviour is to be able to separate out learning and coordination behaviour from simple mimicry. In other words, suppose that farmers simply copy the behaviour of one particular focal individual in the village. Some farmers are better able to observe the behaviour of this individual because they are simply more “sociable”. More sociable farmers are more likely to know more other sociable farmers. This creates correlation between observed behaviour within a village, and the social networks known to each farmer, but the decision to adopt is based on nothing relating to social effects. This is an example of the correlated effects identified by Manski (1993) that was discussed earlier in the section on methodology.

To address this concern we collected data on the behaviour of a key individual in the village that we suppose could act as a focal point for the behaviour of others, namely the farmer perceived by others to be the wealthiest. Not all the such named wealthiest farmers adopted. We control for whether they adopted or not in column (1) of table 11. We find that the behaviour of this individual has no effect in the behaviour of other farmers in the village.³⁴

We would have liked to have tested that if the village elder had adopted, other farmers would adopt. Unfortunately, all village elders in our sample adopt. However, it is possible to identify whether farmer i and the village elder share the same religion. We thus control for whether the village elder adopts and is of the same religion as farmer i . Column (2) shows that this does not have a significant effect on adoption decisions either and does not eliminate our earlier result that information sharing takes place within the village. On this data though, we would not discount that the possibility that mimicry is at least part of the story. However note that the coefficients on the social network are both unchanged. This implies that mimicry may occur but in addition to information sharing.

5 Discussion

The analysis above shows that individual adoption decisions do depend on the adoption choices of network members and that (i) individual networks matter over and above village aggregate measures, (ii) information sharing matters (iii) the relationship between the individual probability of adoption and the number of known adopters

³⁴Similar results were obtained when we controlled for whether the farmer perceived to be the most profitable in the village adopted.

is shaped as an inverse-U.

The fact that information sharing matters is consistent with previous findings in the literature. Both Foster and Rosenzweig (1995) and Conley and Udry (2001) show that farmers learn from each other about new techniques, in particular fertiliser use with new crops. Munshi (1996) also attributes neighborhood effects on technology adoption to social learning. Besley and Case (1998) show that the diffusion process of HYV cotton among a small group of Indian farmers is consistent with a model in which farmers learn from each other about the profitability of the new technology. As suggested by both by the theoretical³⁵ as well as the empirical labour and industrial organization literature social learning and network effects matter in a variety of settings.³⁶

In the context of technology adoption the literature has focussed on two dimensions of learning: learning about the parameters of the new technology and learning about its profitability. In our context the former seems more relevant. Indeed, when asked why they did not adopt most farmers said they did not know how to grow the new crop rather than being uncertain of its profitability. The distinction is, however, irrelevant for the discussion that follows.

Can social learning explain the inverse-U relationship between the individual probability of adoption and the number of known adopters that we observe in the data? The answer depends on whether we analyse the adoption decision at date 0 as part of a dynamic or a static optimisation problem and on whether, in addition to information sharing, networks provide other services whose value depend on the number of adopters within the network.

In a static framework we expect the adoption decision of other network members to affect the individual adoption decision if there are complementarities such that the marginal benefit (cost) of adoption increases (decreases) as the number of adopters increase. Such complementarities might arise, for instance, because farmers expect to learn how to solve problems from other adopters. Complementarities of this type, however, would yield a monotonically increasing, although not necessarily linear, relationship.³⁷ Dynamic considerations, on the other hand, suggest a reason why farmers might be less likely to adopt when they know many adopters. As long as knowledge can be accumulated, farmers have the option of delaying adoption for

³⁵See Banerjee (1992), Caplin and Leahy (1998), Chamley and Gale (1994), Ellison and Fudenberg (1993,1995), Gale (1996), Jovanovic and Nyarko (1996), Kapur (1995), McFadden and Train (1996), Moscarini and Ottaviani (1997).

³⁶Behrman et al (2001) and Munshi and Myaux(2000) analyse network effects in family planning, Bertrand et al (2000) analyse the impact of information sharing on welfare participation, Irwin and Klenow (1994) analyse learning in the semiconductor industry, Thornton and Thompson (1998) study learning spillovers in WWII shipbuilding,

³⁷It could be argued that an increase in the number of adopting farmers drives the price of sunflower down and hence discourages adoption. This explanation, however, is not appropriate in this context for two reasons. First, sunflower is grown for home consumption and commercialisation is currently very difficult (see Section 2). Second, even if farmers sold their output it would be unlikely that the quantity produced were large enough to affect the price.

one period and free ride on the knowledge accumulated by the farmers who have adopted at date 0. The fact that information sharing can lead to strategic delay is widely recognized in the literature³⁸ and does not depend on the precise nature of the shared information. The papers that are more closely related to ours (Besley and Case (1998), Foster and Rosenzweig (1995)) formalise the issue but cannot address it empirically since they have no information on the extension of individual networks.

We argue that static complementarities combined with dynamic incentives to delay can yield a network effect curve as the one we observe in the data, as long as the former prevails when the farmers knows few adopters while the latter dominates when the number of known adopters is high. Such pattern emerges when the complementarity exhibits diminishing returns, i.e. the marginal benefit of adoption increases in the number of adopters but *at a decreasing rate*, and the strategic incentive to delay is stronger when the farmer knows many adopters. This is likely to be the case for the following reason. Delaying adoption is beneficial because the farmer does not cultivate the new crop at date 0 when knowledge about the technology is scarce and profits low. Delaying adoption is costly because since farmer i does not learn at date 0, his profit at date 1 is lower than it would have been had he adopted at date 0. Such cost, however, decreases as the number of farmers who adopt at date 0 increases. In fact, if many farmers adopt at date 0 farmer i 's marginal contribution to the total stock of knowledge and hence on his profits at date 1 becomes negligible.

Alternatively, the observed inverse-U relationship between the individual probability of adoption and the number of known adopters can be explained if we take into account the fact that in addition to information sharing, networks provide other services whose value depend on the number of adopters within the network. One such service is risk-sharing. First, a considerable amount of evidence does suggest that social networks are a very important channel through which individuals insure against shocks (see Fafchamps (1992), Townsend (1994), Udry (1994) and Goldstein (1999)). Second, the effectiveness of the risk-sharing mechanism depends on the number of adopters. Indeed, if many farmers within the same network adopt, diversification and hence the scope for effective insurance within the group are greatly reduced. This in turn can explain why the marginal effect of knowing one more adopter can be decreasing and even negative when most network members are growing the new crop.

Testing between these alternative explanations goes beyond the scope of this paper. An appropriate test would require information on available insurance mechanisms and adoption decisions along the entire diffusion path, which are currently unavailable. We hope to pursue such questions in future research.

³⁸Bardhan and Udry (1999), Besley and Case (1998), Foster and Rosenzweig (1995) analyse the issue in the context of technology adoption in developing countries. For other applications see Caplin and Leahy (1998), Chamley and Gale (1994), Kapur (1995), McFadden and Train (1996).

6 Conclusions

We have presented evidence that “social” effects are an important determinant of individual adoption decisions. The evidence altogether suggest that information sharing is important but that its effect on adoption choice depend on a list of other factors. We have also shown that having a precisely defined social network is crucial for this type of analysis and that the impact of social networks on adoption is highly non-linear. We have argued that non-linearity follows naturally from a theoretical model where network effects derive from information sharing and the adoption decision is analysed as part of a dynamic optimisation problem. We believe that in a variety of contexts where social interactions are relevant, the predicted relationship between individual decisions and group behaviour would be non-linear as well. Estimates of social effects should therefore be backed by a consistent theoretical model that informs about the exact shape of the relationship. Assuming linearity may well lead to misleading empirical results, particularly when information spillovers are the reason why networks matter.

Our data has allowed us to deal with the issues of simultaneity and omitted variables better than what has been feasible with earlier data sets. Yet we are left with many unanswered questions and issues raised by the results. We conclude with a brief discussion of each of these.

First, we find that female headed households are more likely to adopt than males. This is in contrast to some of the earlier literature where female headed household have often been found to be amongst the last to embrace new technologies, possibly because they face greater credit constraints. What is it about females, over and above the other covariates controlled for, that makes them act in such a strikingly different way? Our data allows us to make no further progress in this line of inquiry. What we do know is that there network size is not significantly different from that of males, and there is some evidence that males’ networks are more likely to be church based. It would be useful to see if networks differ in gender composition also. Goldstein (1999) finds for rural Ghana, that women are more likely to insure with women and men insure with a wider and more disparate group. It could be the case that in our sample we have networks mainly comprised of women, who share agricultural information.

Second, the results shed light on the nature of information within the village. We can conceive of a world in which all information is public. In this scenario, it is as if a social planner operates in the village, dividing the villagers between those who should grow and accumulate knowledge about the new crop now and those who would better wait. In such a world, the actual number of adopters known by adopters and non-adopters ought not to differ. This is not what we find, despite the fact that our villages are small. Adopters know significantly more adopters than do non-adopters. Furthermore, the benefits of knowing adopting farmers is greater for more talkative farmers (table 3, column 2). This suggests that there is a distinct private element to information that can only be exploited through one’s own social network. This effect

operates over and above any public information that may be present, as is picked up by the number of other adopters known variable.

Third, endogenous social group formation is a real issue here. The effect of knowing individuals on the basis of some unobservable characteristic, such as the taste for oil, that also determines adoption decisions, has an effect on adoption probabilities (table 5, columns 5 and 6). This has implications for future policy design. For instance, the creation of association groups, which are designed to facilitate the commercialisation of sunflower seeds, could act to widen inequalities within villages if these “high types” decide to join them. Also, any future provisions for group lending may be hampered by the existence of adverse selection across self-selected groups.

In line with other research, one policy intervention supported by our analysis is the provision of insurance against food shocks. Of course, such a policy is not easily implemented. As we have found evidence that mimicry occurs as well as information sharing, another policy intervention that may be appropriate here is the targeting of certain focal individuals in the village, such as village elders, or those that have participated in past NGO projects, although here again an issue of increasing inequality within the village arises.

Finally, it is important to stress that, due to the lack of information on relevant variables, we cannot separate endogenous from exogenous social effects, implying that we cannot quantify the magnitude of the purely endogenous network effect. Social effects operating through these two channels may have different policy implications. Collecting data to address these and the other concerns discussed, ought to be the focus of future research.

7 Appendix: Proofs

Proof of Proposition 1: Assuming all farmers have the same size network, n and suppose that the adoption decision of farmer i is a linear function of his own (single) characteristics, X_i , the adoption decisions of his network, $\sum_{j \neq i}^n a_j$, and the (single) exogenous characteristics of the members of the network, X_j , where u_i is an error term. We derive the OLS estimator of γ if the econometrician, due to lack of data on the characteristics of the social network of i , estimates a model of the form;

$$a_i = \beta X_i + \gamma \left(\sum_{j \neq i}^n a_j \right) + u_i; \quad i = 1, \dots, n \quad (\text{A1})$$

when the true data generating process (DGP) is given by;

$$a_i = \beta X_i + \gamma \left(\sum_{j \neq i}^n a_j \right) + \delta \left(\sum_{j \neq i}^n X_j \right) + u_i; \quad i = 1, \dots, n \quad (\text{A2})$$

where $\gamma \neq 0$ implies the existence of endogenous social effects and $\delta \neq 0$ implies the existence of exogenous social effects. For the time being we abstract away from the possibility of correlated effects so that;

$$\begin{aligned} \text{var}(u_i) &= \sigma_u^2 \\ \text{cov}(u_i, u_j) &= 0 \text{ for all } i \neq j \end{aligned}$$

Writing (A2) in vector notation;

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \beta \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} + \gamma \begin{pmatrix} a_2 + a_3 + \dots + a_n \\ \vdots \\ a_1 + a_2 + \dots + a_{n-1} \end{pmatrix} + \delta \begin{pmatrix} X_2 + X_3 + \dots + X_n \\ \vdots \\ X_1 + X_2 + \dots + X_{n-1} \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

$$\mathbf{a} = \beta \mathbf{X} + \gamma \mathbf{Z} \mathbf{a} + \delta \mathbf{Z} \mathbf{X} + \mathbf{u}$$

where;

$$\mathbf{Z} = \begin{pmatrix} 0 & 1 & \dots & 1 \\ 1 & \ddots & & \vdots \\ \vdots & & \ddots & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix}$$

Hence the reduced form system of the DGP is;

$$\mathbf{a} = [\mathbf{I}_n - \gamma \mathbf{Z}]^{-1} [\beta \mathbf{I}_n + \delta \mathbf{Z}] + [\mathbf{I}_n - \gamma \mathbf{Z}]^{-1} \mathbf{u} \quad (\text{A3})$$

where \mathbf{I}_n is the n -dimensional identity matrix. Note that;

$$[\mathbf{I}_n - \gamma \mathbf{Z}] = \begin{pmatrix} 1 & -\gamma & \dots & -\gamma \\ -\gamma & \ddots & & \vdots \\ \vdots & & \ddots & -\gamma \\ -\gamma & \dots & -\gamma & 1 \end{pmatrix}$$

is symmetric, hence $[\mathbf{I}_n - \gamma \mathbf{Z}]^{-1}$ also takes the symmetric form, with a constant a along the leading diagonal and constant b in all off diagonal entries. Solving the inverse identity equation yields the following pair of equations;

$$\begin{aligned} a - (n-1)\gamma b &= 1 \\ -\gamma a + b - (n-2)\gamma b &= 0 \end{aligned}$$

and solving for (a, b) gives $a = \frac{1-\gamma(n-2)}{1-\gamma^2(n-1)-\gamma(n-2)}$; $b = \frac{\gamma}{1-\gamma^2(n-1)-\gamma(n-2)}$. Hence;

$$[\mathbf{I}_n - \gamma \mathbf{Z}]^{-1} = \alpha(\gamma) \begin{pmatrix} 1 - \gamma(n-2) & \gamma & \dots & \gamma \\ \gamma & \ddots & & \vdots \\ \vdots & & \ddots & \gamma \\ \gamma & \dots & \gamma & 1 - \gamma(n-2) \end{pmatrix}$$

$$\text{where } \alpha(\gamma) = \frac{1}{1 - \gamma^2(n-1) - \gamma(n-2)}$$

Substituting back into the reduced form (A3) gives;

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \alpha(\gamma) \begin{pmatrix} 1 - \gamma(n-2) & \gamma & \cdots & \gamma \\ \gamma & \ddots & & \vdots \\ \vdots & & \ddots & \gamma \\ \gamma & \cdots & \gamma & 1 - \gamma(n-2) \end{pmatrix} \begin{pmatrix} \beta & \delta & \cdots & \delta \\ \delta & \ddots & & \vdots \\ \vdots & & \ddots & \delta \\ \delta & \cdots & \delta & \beta \end{pmatrix} \\ + \alpha(\gamma) \begin{pmatrix} 1 - \gamma(n-2) & \gamma & \cdots & \gamma \\ \gamma & \ddots & & \vdots \\ \vdots & & \ddots & \gamma \\ \gamma & \cdots & \gamma & 1 - \gamma(n-2) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

Hence after multiplying through, the reduced form of the adoption decision of farmer i is given by;

$$a_i = \alpha(\gamma) \left\{ \begin{aligned} & [\beta - \beta\gamma(n-2) + (n-1)\gamma\delta] X_i + (\delta + \gamma\beta) \left(\sum_{j \neq i}^n X_j \right) \\ & + [1 - \gamma(n-2)] u_i + \gamma \left(\sum_{j \neq i}^n u_j \right) \end{aligned} \right\}; \quad i = 1, \dots, n \quad (\text{A4})$$

Estimating (A1) using OLS;

$$\hat{\gamma} = \frac{\text{cov}(a_i, \left(\sum_{j \neq i}^n a_j \right)) \cdot \text{var}(X_i) - \text{cov}(X_i, a_i) \cdot \text{cov}(X_i, \left(\sum_{j \neq i}^n a_j \right))}{\text{var}\left(\left(\sum_{j \neq i}^n a_j\right)\right) \cdot \text{var}(X_i) - \left[\text{cov}(X_i, \left(\sum_{j \neq i}^n a_j\right))\right]^2}$$

We use the reduced form (A4) to calculate the probability limit of each of these terms. Note first that;

$$\left(\sum_{j \neq i}^n a_j \right) = \sum_i^n a_i - a_i$$

$$= \alpha(\gamma) \left\{ \begin{aligned} & [\beta - \beta\gamma(n-2) + (n-1)\gamma\delta] \sum_i^n X_i + (\delta + \gamma\beta) \sum_i^n \left(\sum_{j \neq i}^n X_j \right) \\ & + [1 - \gamma(n-2)] \sum_i^n u_i + \gamma \sum_i^n \left(\sum_{j \neq i}^n u_j \right) \end{aligned} \right\}$$

$$-\alpha(\gamma) \left\{ \begin{array}{l} [\beta - \beta\gamma(n-2) + (n-1)\gamma\delta] X_i + (\delta + \gamma\beta) \left(\sum_{j \neq i}^n X_j \right) \\ + [1 - \gamma(n-2)] u_i + \gamma \left(\sum_{j \neq i}^n u_j \right) \end{array} \right\}$$

The using $\sum_i^n \left(\sum_{j \neq i}^n X_j \right) = (n-1) \sum_i^n X_i$ this expression simplifies to;

$$\left(\sum_{j \neq i}^n a_j \right) = \alpha(\gamma) \left\{ \begin{array}{l} [\beta + (n-1)\gamma\delta + \delta(n-2)] \left(\sum_{j \neq i}^n X_j \right) \\ + (\delta + \gamma\beta)(n-1)X_i + \gamma(n-1)u_i + \left(\sum_{j \neq i}^n u_j \right) \end{array} \right\} \quad (\text{A5})$$

Denoting $\text{var}(X_i) = \sigma_x^2$ for all i , then by substituting (A4) and (A5) into each component in the formula for $\hat{\gamma}$, we obtain;

$$\begin{aligned} \text{plim } \text{cov}(X_i, a_i) &= \alpha(\gamma) [\beta - \beta\gamma(n-2) + (n-1)\gamma\delta] \sigma_x^2 \\ \text{plim } \text{cov}(X_i, \left(\sum_{j \neq i}^n a_j \right)) &= \alpha(\gamma) [(\delta + \gamma\beta)(n-1)] \sigma_x^2 \\ \text{plim } \text{var}\left(\left(\sum_{j \neq i}^n a_j \right)\right) &= \alpha^2(\gamma) \left\{ \left(\begin{array}{l} [\beta - \beta\gamma(n-2) + (n-1)\gamma\delta]^2 (n-1) \\ + (\delta + \gamma\beta)^2 (n-1)^2 \\ + (\gamma^2(n-1)^2 + (n-1)) \sigma_u^2 \end{array} \right) \sigma_x^2 \right\} \\ \text{plim } \text{cov}(a_i, \left(\sum_{j \neq i}^n a_j \right)) &= \alpha^2(\gamma) \left\{ \begin{array}{l} (\delta + \gamma\beta)(n-1) \left[\begin{array}{l} 2\beta + 2\gamma\delta(n-1) \\ + (\delta - \beta\gamma)(n-2) \end{array} \right] \sigma_x^2 \\ (\gamma(n-1) [2 - \gamma(n-2)]) \sigma_u^2 \end{array} \right\} \end{aligned}$$

where recall that $\text{var}(u_i) = \sigma_u^2$ and we are assuming no correlated effects so that $\text{cov}(u_i, u_j) = 0$ for all $i \neq j$. Without loss of generality we set $\sigma_x^2 = 1$. Substituting the above expressions into the probability limit of $\hat{\gamma}$ yields after some manipulation;

$$\text{plim } \hat{\gamma} = \frac{(\delta + \gamma\beta) [\beta + \gamma\delta(n-1) + \delta(n-2)] + [2 - \gamma(n-2)] \sigma_u^2}{[\beta + \gamma\delta(n-1) + \delta(n-2)]^2 + [1 + (n-1)\gamma^2] \sigma_u^2}$$

Hence we can see immediately that in the absence of endogenous and exogenous social effects;

$$\text{plim } \hat{\gamma}|_{\gamma=\delta=0} = 0$$

and in the absence of endogenous social effects only;

$$\text{plim } \hat{\gamma}|_{\gamma=0} = \frac{\delta [\beta + \delta(n-2)]}{[\beta + \delta(n-2)]^2 + \sigma_u^2} > 0$$

The OLS estimate of γ underestimates the true endogenous social effect, i.e. $\hat{\gamma} < \gamma$ when;

$$\begin{aligned} \left[\begin{array}{c} (\delta + \gamma\beta) [\beta + \gamma\delta(n-1) + \delta(n-2)] \\ + [2 - \gamma(n-2)] \sigma_u^2 \end{array} \right] &< \left[\begin{array}{c} \gamma [\beta + \gamma\delta(n-1) + \delta(n-2)]^2 \\ + \gamma [1 + (n-1)\gamma^2] \sigma_u^2 \end{array} \right] \\ \delta [\beta + \gamma\delta(n-1) + \delta(n-2)] [1 - \gamma^2(n-1) - \gamma(n-2)] &< - [1 - \gamma^2(n-1) - \gamma(n-2)] \gamma \sigma_u^2 \end{aligned}$$

which if:

$$\gamma > \frac{1}{n-1} \quad (\text{A6})$$

implies:

$$\delta [\beta + \gamma\delta(n-1) + \delta(n-2)] > -\gamma \sigma_u^2$$

which is always true.

Turning to the case where there exist correlated effects in the data generating process (A2) so that;

$$\begin{aligned} \text{var}(u_i) &= \sigma_u^2 \\ \text{cov}(u_i, u_j) &= \rho \text{ for all } i \neq j \end{aligned}$$

The following expressions now change from the above derivation;

$$\text{plim } \text{var}\left(\sum_{j \neq i}^n a_j\right) = \alpha^2(\gamma) \left\{ \begin{array}{c} \left(\begin{array}{c} [\beta - \beta\gamma(n-2) + (n-1)\gamma\delta]^2 (n-1) \\ + (\delta + \gamma\beta)^2 (n-1)^2 \end{array} \right) \sigma_x^2 \\ + (\gamma^2(n-1)^2 + (n-1)) \sigma_u^2 + (n-1)[2(n-1)\gamma + n-2]\rho \end{array} \right\}$$

$$\text{plim } \text{cov}\left(a_i, \sum_{j \neq i}^n a_j\right) = \alpha^2(\gamma) \left\{ \begin{array}{c} (\delta + \gamma\beta)(n-1) \left[\begin{array}{c} 2\beta + 2\gamma\delta(n-1) \\ + (\delta - \beta\gamma)(n-2) \end{array} \right] \sigma_x^2 \\ (\gamma(n-1)[2 - \gamma(n-2)]) \sigma_u^2 + 2(n-1)[(n-1)\gamma^2 + 1]\rho \end{array} \right\}$$

Inserting these expressions into $\text{plim } \hat{\gamma}$ now yields;

$$\text{plim } \hat{\gamma} = \frac{(\delta + \gamma\beta) [\beta + \gamma\delta(n-1) + \delta(n-2)] + [2 - \gamma(n-2)] \sigma_u^2 + 2(n-1)[(n-1)\gamma^2 + 1]\rho}{[\beta + \gamma\delta(n-1) + \delta(n-2)]^2 + [1 + (n-1)\gamma^2] \sigma_u^2 + (n-1)[2(n-1)\gamma + n-2]\rho}$$

so that in this case even in the absence of any social effects;

$$\text{plim } \hat{\gamma}|_{\gamma=\delta=0} = \frac{\delta(\beta + \delta(n-2)) + 2(n-1)\rho}{(\beta + \delta(n-2))^2 + \sigma_u^2 + (n-1)(n-2)\rho} > 0$$

In this case $\hat{\gamma}$ underestimates the true endogenous social effect if;

$$[\delta(\beta + \gamma\delta(n-1) + \delta(n-2)) + \gamma\sigma_u^2] [1 - \gamma^2(n-1) - \gamma(n-2)] + (n-1)[2 - (n-2)\gamma]\rho < 0$$

a necessary condition for this to hold is, as above,

$$\gamma > \frac{1}{n-1}$$

a sufficient condition is;

$$\gamma > \frac{2}{n-2}. \blacksquare$$

8 Data Appendix

We give definitions of the key variables used in the analysis here. The complete questionnaire is in Gallego-Lizon and Rasul (2001).

Poverty Measures

All the indicators are measures of relative poverty. The asset poverty measure is constructed as follows - a weighted sum of household assets was constructed. The weights were as follows - radio (1), chair(.5), bed (1), water pot (.5), bicycle (2), and jewelry (1). These weights were chosen to reflect local market prices. If a household had a value of weighted assets less than 75% of the average in the sample, they were classified as being “very poor”. If the household had a weighted value of assets above 75% of the average in the sample, the household was classified as “not poor”. All remaining households are “quite poor”. The poverty measure based upon animal ownership was similarly constructed using the following weights - pig (2.5), chicken (1), pigeon (.25), goat (5), canga (.75), duck (1.25), and rabbit (1). The same bands were also used for the poverty measure based upon income. The ranges were experimented with but these widths were chosen because they do allow us to differentiate households between those that suffer from acute food insecurity (low income levels, few possessions or animals), and those that are more food secure.

In addition we asked the enumerators, who are *Movimondo* extension workers, to make a subjective assessment of the individual’s poverty status. Their reported measures correlated quite highly with those we constructed above. For example, for 20% of households the enumerators measure coincided with the simple majority of the constructed measures.

We also constructed a dichotomous poverty measure based upon the households reported consumption (called food poverty). A household was classified as being poor if they do not consume oil nor vegetables at all.

All income variables are self-reported. Income earned was asked by source of income (own farm agriculture, working on other farms, selling drinks, formal sector work, animal production, remittances, pensions, selling stones, other agricultural

work, artisanry, selling wood, hunting, other non-agricultural work). We then aggregated these figures to obtain household income.

Individual Heterogeneity

The measure of having direct knowledge of prices was constructed from the question, “where do you get information from on the price of different cultures?”. The individual could either be uninformed, have their own price information, or receive information from other sources (family, neighbours, NGO extension workers or the radio). The measure of talking to others was similarly constructed.

Social Networks

We asked individuals “how many people do you know that are planting the new cultures?” (none, 1-5, 5-10, 11-20, 21-30, more than 30), and “how many of these belonged to each of the following groups?” (family, neighbours, members of same church, friends for other communities, others). This second question was used also as a check on the correct classification in the first. We asked the same questions regarding expected number of adopters for next planting season to both adopters and non-adopters, and it is these variables that we use to construct our measures of social networks.

Knowledge on Non-agricultural Events

We asked individuals to list the sources, if any, they had obtained information from with regards to two important local events - first, the electrification of the main town in the sample area, Gile, and the rehabilitation of a bridge in the Gile District. The wording of the question was, “how did you find out about the following events?”. If they answered from wife, family, neighbours, friends, church members or the radio, they were classified as not having direct knowledge of these events.³⁹ Otherwise they could have either found out themselves, or did not know about the event.

Social Insurance

The sources of social insurance at the village level were obtained from the village questionnaire. We asked the village elder, “what coping strategies do people use if they have a shortage of food?”. If they replied reduce number of meals eaten, collecting wild fruits or roots, hunting, selling belonging, use savings, work on others’ farms or temporarily migrate, we classify this as having no social insurance. If they replied receive food from family or receive food from neighbours, this classifies this village as having social insurance within the village. If they replied receive food from friends in other villages we classify this as insurance with other villages.

References

- [1] BANDIERA.O AND I.RASUL (2002) “Identification of Network Effects: Does Network Shape Matter?” work in progress.

³⁹Radio is in this category because few individuals own radios themselves, and of those that do, few actually have batteries to operate them.

- [2] BANERJEE. A (1992) "A Simple Model of Herd Behavior" *Quarterly Journal of Economics*; 107(3), pages 797-817
- [3] BARDHAN.P AND C.UDRY (1999) "Development Microeconomics". New York: Oxford University Press.
- [4] BEHRMAN.J.R.,H.P.KOHLER AND S.C.WATKINS (2001) "How Can We Measure the Causal Effects of Social Networks Using Observational Data? Evidence from the Diffusion of Family Planning and AIDS Worries in South Nyanza District, Kenya" *Max Planck Institute for Demographic Research Working Paper* 2001-022
- [5] BENABOU.R (1996) "Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance", *American Economic Review* 86(3): 584-609.
- [6] BERTRAND.M, E.F.P.LUTTMER AND S.MULLAINATHAN, "Network Effects and Welfare Cultures", *Quarterly Journal of Economics* 115(3): 1019-55.
- [7] BESLEY.T AND S.COATE (1992) "Understanding Welfare Stigma: Taxpayer Resentment and Statistical Discrimination", *Journal of Public Economics* 48(2): 165-183.
- [8] BESLEY.T AND A.CASE (1993) "Modelling Technology Adoption in Developing Countries", *American Economic Review* 83(2): 396-402.
- [9] BORJAS.G (1992) "Ethnic Capital and Intergenerational Mobility", *Quarterly Journal of Economics* 107(1): 123-150.
- [10] BORJAS.G (1995) "Ethnicity, Neighborhoods and Human Capital Externalities", *American Economic Review* 85(3): 365-390.
- [11] BROCK.W AND S.DURLAUF (2000) "Interaction Based Models", forthcoming *Handbook of Econometrics, Vol.5.*, Amsterdam, New York and London: Elsevier, North-Holland.
- [12] CAPLIN. A AND J.LEAHY "Miracle on Sixth Avenue: Information Externalities and Search" *Economic Journal*; 108(446), pages 60-74.
- [13] CASE.A AND L.KATZ (1991) "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths", *NBER working paper no.3705*.
- [14] CHAMLEY.C AND D.GALE (1994) "Information Revelation and Strategic Delay in a Model of Investment" *Econometrica*; 62(5), pages 1065-85.
- [15] CONLEY.T AND C.UDRY (2000) Learning About a New Technology: Pineapple in Ghana, Working Paper-Yale University.

- [16] CONLEY.T. AND G.TOPA (2002) "Dynamic Properties of Local Interaction Models", mimeo University of Chicago and NYU.
- [17] DURLAUF.S (1996) "A Theory of Persistent Income Inequality", *Journal of Economic Growth* 1(1): 75-93.
- [18] ELLISON.G AND D.FUDENBERG (1993) Rules of Thumb for Social Learning" *Journal of Political Economy*; 101(4), pages 612-43.
- [19] ELLISON.G AND D.FUDENBERG (1995) "Word-of-Mouth Communication and Social Learning" *Quarterly Journal of Economics*; 110(1), pages 93-125.
- [20] FAFCHAMPS.M (1992) "Solidarity Networks in Preindustrial Societies: Rational Peasants With a Moral Economy", *Economic Development and Cultural Change* 41(1): 147-74.
- [21] FEDER.G, R.E.JUST AND D.ZILBERMAN (1985) "Adoption of Agricultural Innovations in Developing Countries: A Survey", *Economic Development and Cultural Change* 33(2): 255-298.
- [22] FOSTER.A AND M.ROSENZWEIG (1995) "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture", *Journal of Political Economy* 103(6): 1176-1209.
- [23] GALE.D (1996) "What Have We Learned from Social Learning?" *European Economic Review*; 40(3-5), pages 617-28.
- [24] GALLEGO-LIZON.J AND I.RASUL (2001) *Food Security Project in Zambezia Province: Gile and Ile Districts Baseline Survey*, Movimondo Report.
- [25] GLAESER.E, B.SACERDOTE AND J.SCHEINKMAN "Crime and Social Interactions", *Quarterly Journal of Economics* 111(2): 507-548.
- [26] GOLDSTEIN.M (1999) *Chop Time, No Friends, Intrahousehold and Individual Insurance Mechanisms in Southern Ghana*, mimeo, University of California, Berkeley.
- [27] IRWIN.D AND P.J.KLENOW (1994) Learning-by-Doing Spillovers in the Semiconductor Industry" *Journal of Political Economy*; 102(6), pages 1200-1227.
- [28] JENCKS.C AND S.MAYER (1990) "The Social Consequences of Growing Up in a Poor Neighborhood", in L.Lynn and M. McGeary eds. *Inner-City Poverty in the United States*-Washington D.C. National Academy Press.
- [29] JOVANOVIC.B AND Y.NYARKO (1996) "Learning by Doing and the Choice of Technology", *Econometrica*, 64, 6, pp. 1299-1310.

- [30] KAPUR.S (1995) "Technological Diffusion with Social Learning" *Journal of Industrial Economics*; 43(2), pages 173-95.
- [31] KAPTEYN.A AND I.WOITTIEZ (1998) "Social Interactions and Habit Formation in a Model of Female Labour Supply", *Journal of Public Economics* 70(2): 185-205.
- [32] LUCAS.R (1988) "On the Mechanics of Economic Development", *Journal of Monetary Economics* 22(1): 3-42.
- [33] MANSKI.C (1993) "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies* 60(3): 531-542.
- [34] MANSKI.C (1995) *Identification Problems in the Social Sciences* Cambridge: Harvard University Press.
- [35] MCFADDEN.D AND K.TRAIN (1996) "Consumers' Evaluation of New Products: Learning from Self and Others" *Journal of Political Economy*; 104(4), pages 683-703
- [36] MOFFITT.R (1983) "An Economic Model of Welfare Stigma", *American Economic Review* 73(5): 1023-1035.
- [37] MOFFITT.R (2000) *Policy Interventions, Low-Level Equilibria and Social Interactions*, Johns Hopkins University-Working Paper.
- [38] MOSCARINI.G AND M.OTTAVIANI (1997) "Economic Models of Social Learning", in Battigalli,P. Montesano A. and F.Panunzi (eds.) *Decisions, games and markets. Studies in Risk and Uncertainty*. Boston; Dordrecht and London: Kluwer Academic, pages 265-98.
- [39] MUNSHI.K(1996) "Farmers as Econometricians: Social Learning and Technology Diffusion in the Indian Green Revolution", mimeo, University of Pennsylvania
- [40] MUNSHI.K AND J.MYLAUX (2000) *Social Effects in Demographic Transition: Evidence from Matlab, Bangladesh*, University of Pennsylvania-Working Paper.
- [41] ROMER.P (1986) "Increasing Returns and Long Run Growth", *Journal of Political Economy* 94(5): 1002-1037.
- [42] SILVERMAN.B.W (1986) *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.
- [43] THORNTON.R AND P.THOMPSON (1999) Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding" *American Economic Review*; 91(5): 1350-68.

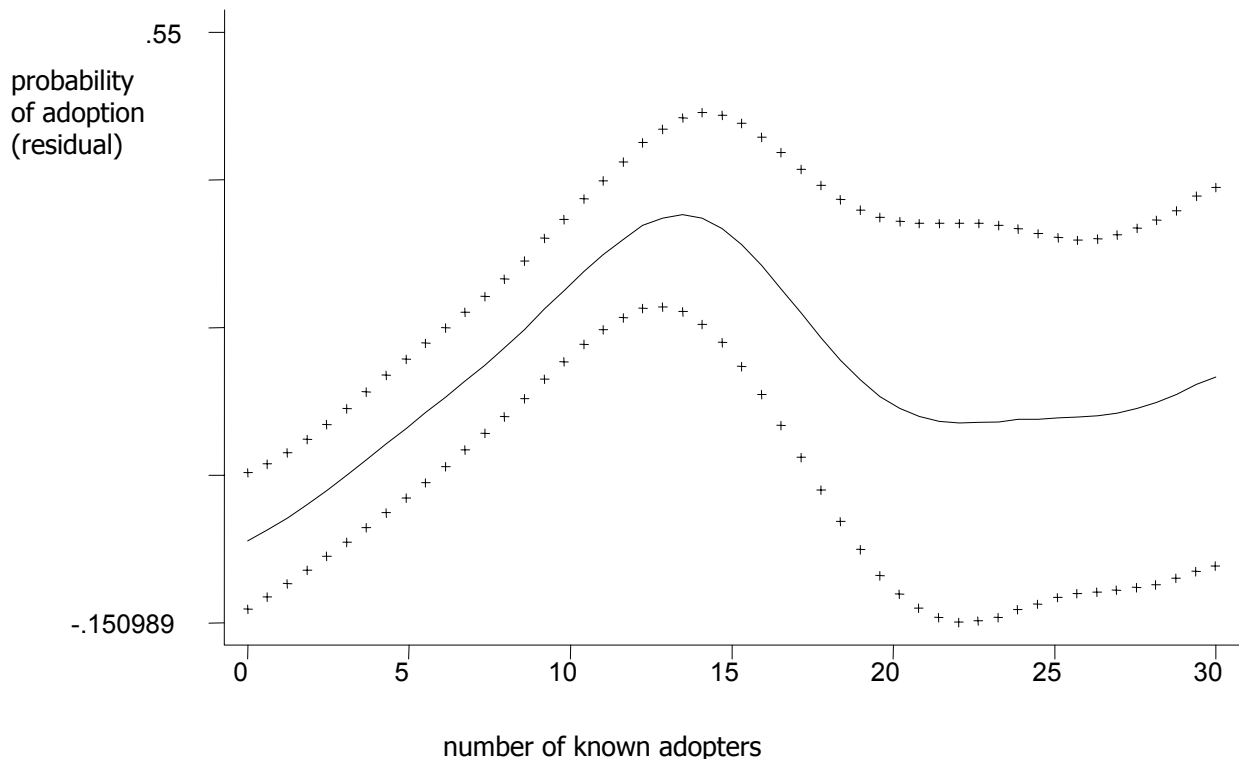
- [44] TOWNSEND.R (1994) "Risk and Insurance in Village India", *Econometrica* 62(3): 539-91.
- [45] UDRY.C (1994) "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria", *Review of Economic Studies* 61(3): 495-526.
- [46] WOITTEZ.I AND A.KAPTEYN (1998) "Social Interactions and Habit Formation in a Model of Female Labour Supply" *Journal of Public Economics*; 70(2): 185-205.



Study Area

Full Area:	801,600 sq km (309,500 sq mi)
Population:	15.7 million
Capital city:	Maputo (pop 1.3 million)
People:	African (99%, including Shangaan, Chokwe, Manyika, Sena and Makua)
Language:	Portuguese (official), indigenous languages
Religion:	Indigenous beliefs (50%), Christian (30%), Muslim (20%)
GDP:	US\$12.2 billion
GDP per head:	US\$670
Growth:	3%
Inflation:	22%
Major industries:	Food, beverages, textiles, chemicals, petroleum

FIGURE 1. NETWORK EFFECTS: NON-PARAMETRIC ESTIMATES



Notes: The variable on the vertical axis is the individual adoption decision (no=0, yes=1) controlling for the variation explained by individual and village characteristics. In each case, the kernel function is Gaussian. Dotted lines are bootstrapped 95% confidence bands, using 1000 replications. Bandwidths in each case were chosen according to;

$$b = 1.06 \min(\sigma, 0.75IQR) N^{-\frac{1}{5}}$$

where σ is the standard deviation of the density, IQR is the interquartile range, and N is the sample size (Silverman (1986)). We used 50 grid points, and the results proved to be robust to the choice of kernel, bandwidth and number of grid points.

Table 1: Descriptive Statistics (standard deviations in parentheses)

1A. Reported Reasons For Adopting (or not)

WHY ADOPTERS ADOPTED

	0.67
I thought it would be remunerative	(0.47)
The NGO convinced me it would be remunerative	0.36 (0.48)
Others convinced me it would be remunerative*	0.15 (0.35)
Existing crops are not remunerative	0.03 (0.17)
	0.72
I want to consume the new crop	(0.45)

* "Others" include one or more of the following: family member, friend, church member, village elder, wife. Sample is 204 farmers in villages where Movimondo operates.

WHY NON-ADOPTERS DID NOT ADOPT

I thought it would not be remunerative/no market for this crop	0.31 (0.47)
I did not know the production techniques	0.48 (0.5)
Somebody else convinced me it would not be remunerative	0.02 (0.14)
Existing crops are remunerative	0.03 (0.17)
No land available	0.17 (0.38)
Other	0.13 (0.33)

1B: Direct Evidence on Network Effects

WOULD YOU BE MORE WILLING TO ADOPT IF..

one family member would also adopt	.73 (.45)
one of your neighbours/friends would also adopt	.16 (.37)
one church member would also adopt	.40 (.50)
your wife would also adopt	.32 (.47)
one of your friends from another village would also adopt	.05 (.22)

Sample - 37 farmers in 2 villages where seeds have not yet been offered.

Table 1 Continued: Descriptive Statistics

Standard Error Reported in Parentheses

	Adopters	Non-Adopters	Difference	Test : p-value
DEMOGRAPHIC DATA				
Female Headed Households	.206 (.039)	.093 (.029)	.113 (.049)	.025
Can Write	.645 (.046)	.567 (.050)	.078 (.068)	.2554
Age	41.766 (1.210)	39.021 (1.35)	.246 (1.81)	.1318
Months of Food Security	9.08 (.165)	8.89 (.157)	.198 (.228)	.3876
Log(income) (Mts)	12.30 (.259)	12.75 (.163)	-.442 (.305)	.1505
ASSET POVERTY (proportion in each group)				
Very Poor	22.43	21.65		
Quite Poor	49.53	49.48		
Not Poor	28.04	28.87		
INDIVIDUAL HETEROGENEITY MEASURES				
Past Participation in NGO Projects	.224 (.040)	.113 (.032)	.111 (.052)	.0359
Direct Knowledge of Agricultural Prices	.654 (.046)	.515 (.051)	.139 (.068)	.0443
Talks to Others About Agriculture	.850 (.034)	.711 (.046)	.139 (.057)	.0158
SOCIAL NETWORKS				
Number of Other Known Adopters	11.96 (1.24)	6.97 (.924)	4.99 (1.55)	.0015
Know Family Members Adopting	.682 (.045)	.402 (.050)	.280 (.067)	.0001
Know Neighbours Adopting	.794 (.039)	.464 (.051)	.330 (.064)	.0000
Know Church Members Adopting	.290 (.044)	.155 (.037)	.135 (.057)	.0211
Number of Family Members Known Adopting	3.16 (.370)	2.35 (.40)	.808 (.542)	.1374
Number of Neighbours Known Adopting	3.70 (.386)	1.77 (.262)	1.93 (.467)	.0001
Number of Church Members Known Adopting	2.43 (.496)	.804 (.247)	1.63 (.554)	.0039
SOCIAL NETWORKS BY GENDER OF HH HEAD				
	Male	Female	Difference	Test : p-value
Know Family Members Adopting	.526 (.038)	.677 (.084)	-.151 (.092)	.1187
Know Neighbours Adopting	.618 (.037)	.742 (.079)	-.123 (.087)	.1880
Know Church Members Adopting	.225 (.032)	.226 (.075)	-.0004 (.081)	.9963
Number of Family Members Known Adopting	2.74 (.306)	2.97 (.536)	-.228 (.617)	.7133
Number of Neighbours Known Adopting	2.61 (.266)	3.74 (.632)	-1.13 (.686)	.1074
Number of Church Members Known Adopting	1.70 (.325)	1.42 (.610)	.280 (.691)	.6870

Exchange rate in June 2001 was US\$1 = 21235 Meticash (Mts). Test of difference based on test of proportions or test of means, not imposing the restriction that variances are equal. The null hypothesis is that the proportion/means are equal, against a two-sided alternative. P-values are reported at the 5% significance level. There are 35 female headed households in the sample.

Table 1 continued: Village Descriptives

	Chiraco	Inlepa	Invali	Jajo	Khayane	Mania-Naholoco	Morrua	Namucarrau	Turbina
Total Households	20	19	34	18	32	20	27	27	14
Male Headed	17	10	31	15	27	18	20	24	11
Actual Aggregate Adoption Rate	.29	.24			.12	.21	.58		.25
Travel Time To Nearest Road on Foot (minutes)	50	2	60	30	30	10	60	180	60
Median Oil Consumption (days per month)	4	11	12	8	2	4	4	0	0
Mean Oil Consumption (days per month)	6	15	13	13	6	9	4	2	5
Other NGO operates in village	⊗	⊗	⊗		⊗	⊗		⊗	
Food Market	⊗	⊗		⊗	⊗	⊗	⊗		
Well		⊗							
School	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗	
Health Post	⊗	⊗		⊗					
Insurance Within Own Village	⊗		⊗	⊗	⊗			⊗	⊗
Insurance With Other Villages		⊗	⊗	⊗	⊗			⊗	⊗

Table 2: Baseline Determinants of Adoption

Linear Probability Model

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise

T-statistics reported in parentheses

	(1)	(2)	(3)	(4)	(5)
Female Headed Household	.370 (4.01)	.498 (4.29)	.476 (4.10)	.367 (3.71)	.387 (3.96)
Can Write	.245 (3.42)	.397 (4.01)	.382 (3.74)	.210 (2.64)	.187 (2.39)
Age	.027 (1.79)	.041 (1.77)	.039 (1.73)	.025 (1.61)	.023 (1.55)
Age Squared x	-.208 (1.24)	-.348 (1.33)	-.343 (1.32)	-.208 (1.23)	-.200 (1.20)
Asset Poverty (Quite Poor)	-.053 (0.59)	.007 (.06)	-.003 (.03)	.0006 (.01)	-.015 (.16)
Asset Poverty (Not poor)	-.132 (1.28)	.0002 (.00)	.012 (.09)	.025 (.23)	-.016 (.14)
Months of Food Security	.027 (1.33)	.019 (.65)	.044 (1.53)	.049 (1.70)	.051 (1.79)
Past Participation in NGO Projects	.232 (2.78)	.049 (.65)	.030 (.22)	.389 (3.66)	.370 (3.58)
Direct Knowledge of Agricultural Prices	.199 (2.23)	.10 (1.03)	.065 (.63)	.058 (.75)	.066 (.88)
Talks to Others About Agriculture	.297 (3.26)	.245 (1.94)	.201 (1.52)	.295 (2.37)	.301 (2.54)
Village Adoption Rate		.522 (1.76)	.482 (1.65)		
Number of Other Known Adopters			.011 (1.64)	.056 (4.19)	
Number of Other Known Adopters Squared x				-.181 (3.85)	
Know 1-9 Adopters					.185 (2.18)
Know 10-16 Adopters					.503 (4.81)
Know More Than 16 Adopters					.203 (1.60)
Village Fixed Effects	No	No	No	Yes	Yes
R-squared	.1950	.2510	.2708	.3366	.3486
Test 1 (p-value reported)					.000
Tests 2 (p-value reported)					.013
Observations	200	112	112	183	183

All regressions are robustly estimated using Eicher-White variance covariance matrices. The omitted category for the poverty measure is "very poor". Complete variable definitions are given in the data appendix. Test 1 in column (5) is for the equality of the coefficients on knowing 1-9 and 10-16 adopters. Test 2 in column (5) is for the equality of the coefficients on knowing 10-16 and more than 16 adopters.

Table 3: Social Networks and Information Sharing

Linear Probability Model

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise

T-statistics reported in parentheses

	(1)	(2)	(3)	(4)
Female Headed Household	.367 (3.71)	.343 (3.57)	.375 (3.77)	.352 (3.54)
Can Write	.210 (2.64)	.203 (2.58)	.212 (2.78)	.203 (2.68)
Age	.025 (1.61)	.025 (1.60)	.028 (1.80)	.030 (1.91)
Age Squared x	-.208 (1.23)	-.211 (1.26)	-.236 (1.41)	-.253 (1.52)
Asset Poverty (Quite Poor)	.0006 (.01)	-.007 (.08)	.016 (.17)	.017 (.19)
Asset Poverty (Not poor)	.025 (.23)	-.008 (.08)	-.008 (.08)	-.020 (.18)
Months of Food Security	.049 (1.70)	.051 (1.80)	.060 (2.20)	.060 (2.23)
Past Participation in NGO Projects	.389 (3.66)	.382 (3.58)	.351 (3.58)	.331 (3.34)
Direct Knowledge of Agricultural Prices	.058 (.75)	.055 (.73)	.079 (1.08)	.073 (.99)
Talks to Others About Agriculture	.295 (2.37)	.149 (1.13)	.345 (2.86)	.108 (.67)
Number of Other Known Adopters	.056 (4.19)	.043 (2.61)		
Number of Other Known Adopters Squared x	-.181 (3.85)	-.212 (4.42)		
Talks to Others About Agriculture x Number of Other Known Adopters		.025 (1.72)		
Knows at Least One Adopter			.302 (3.77)	.088 (.59)
Knows at Least One Adopter x Talks to Others About Agriculture				.318 (1.74)
Village Fixed Effects	Yes	Yes	Yes	Yes
R-squared	.3366	.3510	.2771	.2895
Observations	183	183	200	200

Table 4: Heterogeneity Measures

Linear Probability Model

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise

T-statistics reported in parentheses

	(1)	(2)	(3)	(4)	(5)	(6)
Number of Other Known Adopters	.016 (3.05)	.016 (3.07)	.012 (2.40)			
Knows at Least One Adopter				.333 (3.84)	.331 (4.03)	.302 (3.77)
Exogenous Individual Characteristics	No	Yes	Yes	No	Yes	Yes
Individual Heterogeneity Measures	No	No	Yes	No	No	Yes
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.0916	.1960	.2845	.0916	.1774	.2771
Observations	183	183	183	200	200	200

Robust standard errors in parentheses. Exogenous individual characteristics are gender, can write, age, age squared, asset poverty, months of food security. Individual heterogeneity measures are whether participated in any other NGO projects in the past, has obtained knowledge of local agricultural prices by oneself, and whether the individual talks to others about agriculture.

Table 5: Endogenous Social Networks

Linear Probability Model

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise

T-statistics reported in parentheses

	(1)	(2)	(3)	(4)	(5)	(6)
Female Headed Household	.332 (3.37)	.315 (2.91)	.438 (2.87)	.367 (3.71)	.207 (2.89)	.240 (2.72)
Can Write	.216 (2.91)	.189 (2.25)	.165 (1.86)	.210 (2.64)	.077 (1.18)	.222 (2.84)
Age	.028 (1.94)	.024 (1.50)	.022 (1.35)	.025 (1.61)	.009 (0.71)	.016 (1.10)
Age Squared x	-.245 (1.51)	.209 (1.19)	-.0002 (1.06)	-.208 (1.23)	-.00 (.62)	-.010 (.63)
Asset Poverty (Quite Poor)	-.031 (.35)	.002 (.02)	.013 (0.14)	.0006 (.01)	-.022 (.33)	-.096 (1.10)
Asset Poverty (Not poor)	-.056 (.52)	.032 (.26)	.072 (.59)	.025 (.23)	.085 (1.07)	-.17 (1.76)
Months of Food Security	.076 (2.97)	.043 (1.55)	.039 (1.36)	.049 (1.70)	.036 (1.45)	.067 (2.95)
Past Participation in NGO Projects	.333 (3.54)	.416 (3.81)	.405 (3.69)	.389 (3.66)	.333 (3.46)	.283 (3.10)
Direct Knowledge of Agricultural Prices	.074 (1.05)	.082 (1.06)	.081 (1.02)	.058 (.75)	.088 (1.40)	.228 (3.39)
Talks to Others About Agriculture	.293 (2.39)	.303 (2.37)	.333 (2.70)	.295 (2.37)	.250 (2.02)	.209 (2.25)
Know Family Members Adopting	.163 (2.04)					
Know Neighbours Adopting	.275 (3.65)					
Know Church Members Adopting	.046 (0.54)					
Number of Family Members Known Adopting		.059 (1.98)	.072 (1.75)			
Number of Family Members Known Adopting Squared		-.604 (2.36)	-.623 (2.43)			
Number of Neighbours Known Adopting		.075 (3.18)	.058 (1.95)			
Number of Neighbours Known Adopting Squared		-.298 (3.00)	-.304 (2.95)			
Number of Church Members Known Adopting		.005 (.15)	-.054 (1.41)			
Number of Church Members Known Adopting Squared		.048 (.23)	-.079 (.35)			
Male x Number of Family Members Known Adopting			-.014 (.43)			
Male x Number of Neighbours Known Adopting			.022 (.78)			
Male x Number of Church Members Known Adopting			.086 (2.40)			
Number of Other Known Adopters				.056 (4.19)	.037 (2.99)	.035 (2.71)
Number of Other Known Adopters Squared x				-.181 (3.85)	-.106 (2.36)	-.126 (2.42)
Average Number of Adopters Known in Adoption Group					.121 (12.31)	.039 (4.35)
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	No
R-squared	.3601	.3574	.3750	.3366	.5989	.3254
Test 1 (p-value)		.6943	.8095			
Observations	204	183	183	183	183	183

Test 1 tests the equality of the coefficients of number of family members and number of friends.

Table 6: Probit Estimates

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise
 T-statistics reported in parentheses

	(1)	(2)	(3)	(4)
Female Headed Household	.472 (3.87)	.485 (4.13)	.450 (3.59)	.418 (3.35)
Can Write	.324 (2.90)	.282 (2.45)	.313 (2.78)	.310 (2.66)
Age	.041 (1.96)	.042 (1.98)	.041 (1.96)	.037 (1.75)
Age Squared x	-.033 (1.47)	-.037 (1.57)	-.035 (1.52)	-.031 (1.35)
Asset Poverty (Quite Poor)	.025 (.20)	-.031 (.25)	.012 (0.10)	.023 (.18)
Asset Poverty (Not poor)	.066 (.43)	.005 (.04)	.043 (.27)	.034 (.21)
Months of Food Security	.074 (1.61)	.083 (1.79)	.080 (1.71)	.067 (1.44)
Past Participation in NGO Projects	.449 (3.71)	.418 (3.48)	.444 (3.66)	.458 (3.70)
Direct Knowledge of Agricultural Prices	.099 (1.01)	.121 (1.20)	.102 (1.04)	.130 (1.29)
Talks to Others About Agriculture	.378 (2.82)	.416 (2.98)	.205 (1.14)	.398 (2.91)
Number of Other Known Adopters	.088 (4.43)		.074 (3.46)	(2.44)
Number of Other Known Adopters Squared x	-.296 (3.87)		-.338 (4.16)	(2.69)
Know 1-9 Adopters		.291 (2.50)		
Know 10-16 Adopters		.604 (4.61)		
Know More Than 16 Adopters		.298 (1.92)		
Talks to Others About Agriculture* Number of Other Known Adopters			.030 (1.74)	
Number of Family Members Known Adopting				.100 (2.44)
Number of Family Members Known Adopting Squared				-.940 (2.69)
Number of Neighbours Known Adopting				.105 (3.46)
Number of Neighbours Known Adopting Squared				-.445 (2.57)
Number of Church Members Known Adopting				-.017 (.36)
Number of Church Members Known Adopting Squared				.003 (.75)
Village Fixed Effects	Yes	Yes	Yes	Yes
Model P-value	.0000	.0000	.0000	
Observations	183	183	183	

Table 7: Instrumental Variables Regression

T-statistics reported in parentheses

First Stage Regressions (OLS)

Dependent Variable: Number of Other Adopters Known

Source of Information for Following Projects;		(1)
Electrification of Gile	Yourself	.79 (.35)
	Others	4.01 (1.91)
Number of Sources of Information on Electrification of Gile		-3.94 (3.41)
R-squared		.3236
F-test of underidentification (p-value)		.0050
Observations		183

Second Stage Regressions (Linear Probability Model)

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise

	(1)
Female Headed Household	.289 (2.06)
Can Write	.212 (2.12)
Age	.029 (1.61)
Age Squared x	-.028 (1.35)
Asset Poverty (Quite Poor)	-.095 (.74)
Asset Poverty (Not poor)	-.072 (.49)
Months of Food Security	.131 (1.92)
Past Participation in NGO Projects	.279 (2.10)
Direct Knowledge of Agricultural Prices	.035 (.37)
Talks to Others About Agriculture	.217 (1.42)
Number of Other Known Adopters	.056 (1.87)
Village Fixed Effects	Yes
Test of Overidentification (p-value)	.6913
Observations	183

Table 8: Omitted Village Characteristics

T-statistics reported in parentheses

Number of Other Known Adopters	.054 (4.19)
Number of Other Known Adopters Squared x	-.178 (3.85)
Other NGO Projects Operate in Village	-.203 (1.99)
Distance from Main Road	.004 (3.47)
External Insurance Available	-.018 (.18)
Food Market	.162 (.99)
Well	.139 (.86)
School	-.092 (.50)
Health Post	-.047 (.48)
Joint Test of Village Variables (p-value reported)	.0118
R-squared	.3355
Observations	189

Table 9: Alternative Poverty Measures

Dependent Variable = 1 if household head decides to adopt sunflower, 0 otherwise

T-statistics reported in parentheses

	Animal Ownership	Food Consumption	Log (income)	Enumerator's Evaluation
Number of Other Known Adopters	.058 (2.08)	.055 (3.82)	.054 (4.06)	.056 (4.25)
Number of Other Known Adopters Squared x	-.191 (3.85)	-.179 (3.85)	-.176 (3.73)	-.182 (3.84)
Village Fixed Effects	Yes	Yes	Yes	Yes
R-squared	.3506	.3363	.3557	.3391
Observations	183	183	183	183

Table 10: Imitation Effects

T-statistics reported in parentheses

	(1)	(2)
Number of Other Known Adopters	.042 (2.61)	.043 (2.76)
Number of Other Known Adopters Squared x	-.211 (4.42)	-.220 (4.78)
Talks to Others About Agriculture x Number of Other Known Adopters	.024 (1.72)	.025 (1.82)
Richest Farmer in Village Adopts	.166 (1.03)	
Village Elder is of Same Religion and Adopts		.128 (1.40)
Village Fixed Effects	Yes	Yes
R-squared	.3510	.3599
Observations	183	183