Social Effects in Growing Bananas: Evidence from a Tanzanian Village.

June 2006

Katleen Van den Broeck

Abstract: We analyse the effects of information flows in a small Tanzanian village. Firstly, we examine social effects in technique adoption. Secondly, we test whether agricultural information flows give rise to social learning effects in banana cultivation. Therefore we study how lagged average yield in a reference group a farmer belongs to, affects that individual farmer's harvest. For our analysis we have data on three social groups, namely kinship members, neighbours and social support group members. We find strong exogenous social effects of group education on individual farmers' output in kinship and social support groups and weak public effects of education in neighbourhoods. But only within kinship related groups we find the endogenous effects that produce the social multiplier in banana cultivation.

1. Introduction

In Kagera, in North-western Tanzania, population pressure is high, land becomes scarce and soil fertility decreases. In the long run these evolutions may threaten food security. The local staple food is the Eastern African Highland banana, which becomes ever more prone to diseases and pests. Therefore, NGOs and farmer extension centres introduced new technologies in the region. Farmers can decide to plant new banana varieties, originating from other parts of Eastern Africa or from Latin America, or to adopt techniques that mitigate the negative effects of diseases and weevil attacks. However, the farmer extension officers believe that effects on total banana output will be minor because only few farmers come to learn techniques. But this is without taking into account social learning and social multiplier effects.

A broad range of literature on the adoption of a new technology exists[1]. Before adoption, a farmer has to become aware of a new technology, which can occur via different sources. A farmer can directly learn about the technology through extension services or other forms of communication, e.g. pamphlets, books, radio. These farmers are the innovators or the experimenting farmers in the village and it are these farmers only that are observed by the extension officers. But other farmers learn about a new technology through the innovating farmers. Hence there are externalities attached to the few farmers that learn agricultural methods directly from extension officers. With respect to actual adoption, it is often found that the probability of adoption is influenced by individual characteristics of the farmer such as his human capital, the degree of risk aversion and farm size. Additional to a farmer's own characteristics, his behaviour can be influenced by other farmers' actual behaviour or their attitude towards the new technology. Extensive work is done on this type of social interactions in technology adoption (Case, 1992; Ellison and Fudenberg, 1993, 1995; Foster and Rosenzweig, 1995; Gale, 1996; Bardhan and Udry, 1999; Udry and Conley, 2001; Bandiera and Rasul, 2001).

The ambition of technique diffusion is to increase productivity. In the case of banana growing in Kagera technique adoption is to increase banana output per unit of (scarce) land. If farmers learn techniques from other farmers in the right way (exact

transmission of how the technique has to be applied), the adoption of techniques does not only affect output of the innovating farmer but also output of the "learning" farmer. However, the problem may be that farmers are not as good teachers as the extension officers are which can reduce the social multiplying effect. Thus our analysis will be twofold: besides social effects in mere technique adoption we will analyse social effects in banana output, which is an indirect measure of farmers' behaviour and the final aim of technique diffusion. Additional to a farmer's characteristics, we analyse how the output of a reference group affects output of an individual farmer belonging to that group. We assume and provide evidence that the underlying channel for this process is information flows (e.g. the transmission of productivity enhancing techniques).

Analysing social interactions is not straightforward. Positive effects of group behaviour on individual behaviour can easily be misinterpreted as social effects, while in fact they are due to characteristics common to all members in a reference group. Severe identification problems exist while analysing contemporaneous behaviour. This problem was defined by Manski (1993, 2000) as the reflection problem. Group behaviour influences individual behaviour but each individual affects mean group behaviour so there is a simultaneity bias. Manski and others (e.g. Brock and Durlauf, 2001) have investigated alternative models that mitigate the simultaneity in the model.

Amongst other solutions, they propose to have individual behaviour vary with lagged rather than with contemporaneous group behaviour. A precondition is that the researcher a priori knows the appropriate lag length. With respect to the lag length in growing bananas we believe we have data that cover the appropriate time lag. We have data on a small village in Tanzania. The survey covered the course of one year, but some questions on agricultural output were asked retrospectively about the year before the survey. Since a banana tree needs approximately one year to reach the flowering stage, we assume that the one year lag is appropriate to measure possible social interaction effects. Another alternative Manski and others propose is to use a non-linear model, which also presumes knowing the correct non-linear function. Or one could use another feature of group behaviour, such as the median instead of the mean, but again one has to know a priori the relevant feature. And the last alternative

[1] A survey of the literature on adoption of agricultural innovations in developing countries can be found in Feder, Just and Zilberman, 1985; Besley and Case, 1993.

they offer is to use instrumental variables that directly affect outcome of some but not all group members.

Another problem researchers encounter when using datasets that were not specifically collected to analyse social effects, is the identification of the reference group. Manski argues that not knowing the exact group composition seriously worsens the reflection problem. Because of data constraints many authors use geographical reference group boundaries such as the whole village (Foster and Rosenzweig, 1995), district (Case, 1992) or ethnicity (Borjas, 1992, 1995) as the reference group. These expected reference groups are usually very large while the relevant group may be at a much smaller level[2]. From our survey[3] we found that the persons inside the village a farmer learns techniques from, or asks agricultural advice from (representing a farmer's agricultural information network), are significantly more often kinship related, social support network members, clan members or living relatively close together, compared to the households that do not have a learning or advice link. Therefore, we believe these groups are good proxies for the correct agricultural information group. For these groups we have full information on all members living in the same village (there was no sampling procedure; all the households in the village were included in the survey).

Having information on all the group members gives us two advantages to solve the simultaneity problem. Firstly, we can exclude the behaviour of the farmer himself from his average lagged group behaviour. Secondly, within each group we can identify the (relatively) most productive farmers. Since we assume the most productive farmers will be teachers rather than learners, including them could bias our results. The existence of a direction to the learning effects was acknowledged by Case (1992).

For the analysis, we proceed as follows. Since we assume information flows between farmers to be the underlying cause for social effects in banana growing we define some social groups that can serve as proxies for a farmer's agricultural information network by analysing learning and advice links in the village. Next we use these

[2] Brock and Durlauf (2001) describe the difference between global and local interactions. In the case of global interactions each individual assigns an identical weight to the behaviour of every other member of the population. In the case of local interactions the model assumes that each agent interacts directly with only a finite number of others in the population.

[3] The survey contains questions about agricultural advisers such as "if you encountered a serious problem with any of your crops whom would you turn to for advice". There is also a section on banana growing techniques known by the farmer (household head) and whom he learned them from.

groups firstly to find how group behaviour with respect to technique adoption affects individual technique adoption. Secondly we test whether the outcome effect of information flows is the same in each of these groups by analysing the effect of lagged average group output on individual banana output.

In the next section we will present an analytical framework for our analysis. The framework for social effects in technique adoption is mainly drawn from Berger (1985) and Udry and Conley (2001). Our social learning test in growing bananas is based on Manski (1993). In section three we continue with the description of the data with respect to learning and advice seeking behaviour of the farmers in a small village in Tanzania. We analyse information flows and define proxies for a farmer's agricultural information network. Section four is the empirical part, where we test for social interactions between farmers. We look at how adoption behaviour is related between farmers of the same social group and whether individual banana output is affected by the lagged average banana yield in a reference group (which was called by Manski the endogenous social effect). We correct for individual characteristics and exogenous group characteristics such as education level and age composition of the group (exogenous social effects).

In the survey village, we find the strongest endogenous social effects within kinship related groups. We do not find this type of social effects within neighbours or social support groups once corrected for exogenous characteristics. Not correcting for exogenous effects leads to the false conclusion of the existence of an endogenous effect in social support networks. It suggests that these groups are probably formed on the basis of characteristics that drive not only group formation, but output of the group members in the same way. Complementary to the private effects of some years of education, we do find strong exogenous social effects of education levels of the group members in social support and kinship related groups, and weak educational effects in neighbourhoods. But only within kinship related groups we find the endogenous effects that produce the social multiplier effect, which may be caused by a better transmission of information.

2. Analytical framework

Section two will provide on the one hand a framework for farmers' updating of beliefs with respect to the expected benefit of a technique, based on the experiences of other farmers. On the other hand it will tackle the problems and econometric solutions for analysing social learning based on inter-farmer information flows.

2.1. Updating of beliefs through other farmers' experiences

We use the theoretical framework on Bayesian updating of beliefs provided in chapter four of Berger (1985) and the target input model (Udry and Conley, 2001; Foster and Rosenzweig, 1995) as the basis for our empirical test of social interactions in technique adoption. In target input models the optimal level of inputs is unknown and stochastic. The target input model starts from the point that an individual has to make a quantitative decision about an action to undertake, more specifically about the right amount of input when applying a new technology. A Bayesian framework is used where farmer *i* has some prior beliefs about the right amount of input and he updates his beliefs each period. New information is revealed by experimenting himself or by the experiments of other farmers. We use the same Bayesian framework to analyse farmers' updating of beliefs with respect to actual benefits of technique adoption[4].

Assume farmer *i* has prior beliefs about the benefit (production gain minus effort) of adopting a technique. In the first period he believes the production gain of applying the technique is not worth the effort[5]. Each period more information is revealed about the true benefits by other farmers, who do apply the technique, and farmer *i* updates his beliefs using this newly revealed information. When farmer *i*'s expectations turn positive, he adopts the technique. We will first assume that information can be transmitted without any noise and flows between all farmers in the village in the same way. Later we will introduce noise in the process, the variance of which may be dependent on the relationship between two farmers.

[4] Besley and Case (1994) also analysed the framework for updating expectations on profitability of HYV cotton.

[5] There is a question in the survey that confirms this assumption. When farmers were asked about the reasons of not applying a technique they knew, many answered that the technique was too difficult, not yet necessary on their farm or too costly.

We assume that the true benefit of the technique X, which is assumed to be identically and independently distributed across time and farmers, is a normally distributed function, of which all farmers know the variance σ^2 -since they know the banana production function and no additional risk is attached to the techniques- but not the mean, B. At *t-1* farmer *i* has prior beliefs on the benefits, B_{t-1} (production gain minus effort) of applying a technique. Suppose the prior B_{t-1} is a normally distributed function with unknown mean β_{t-1} and known variance τ_{t-1}^2. So in period *t-1* farmer *i* believes that the benefit of applying the technique is β_{t-1}, which we assume to be negative and he does not adopt the technique. Other farmers with other (positive) priors do apply the technique. Suppose farmer *i* can perfectly observe the production gain on other farmers' fields and can learn about the effort put into it without noise in period *t*. Farmer *i* updates his beliefs about the benefits of the technique taking into account his own prior beliefs and the information revealed by the experimenting farmers ($\bar{x}_t$ is the average benefit of the n experimenting farmers). The more convinced he is about his own prior beliefs, the less weight the experimental outcomes will have in his updating. The posterior function $f(B|X)$ will also be normally distributed with mean $\beta_t(x)$ and variance ρ_t^{-1}. If n farmers are experimenting and the information from their experiments reaches farmer *i* in the same way, then it holds that (Berger, 1985; Foster and Rosenzweig, 1995; Udry and Conley, 2001)

$$\beta_t(x) = \frac{\frac{\sigma^2}{n}}{\tau_{t-1}^2 + \frac{\sigma^2}{n}} \beta_{t-1} + \frac{\tau_{t-1}^2}{\tau_{t-1}^2 + \frac{\sigma^2}{n}} \bar{x}_t \qquad (1)$$

$$\rho_t^{-1} = \left(\frac{n}{\sigma^2} + \frac{1}{\tau_{t-1}^2} \right)^{-1} \qquad (2)$$

When $\beta_t(x)$ becomes positive farmer *i* starts applying the technique.

In the presented case there is no noise in the transmission of information and farmer *i* learns from all n experimenting farmers in the same way. However, the assumption of

perfect and full information is rather unrealistic[6]. Therefore we will adapt (1) and (2) to hold for imperfect information and for its belief updating effect to be dependent on the relationship between farmers.

Thus information does not reach farmer *i* perfectly, but with some noise. Instead of observing x_t he observes $x_t + u_t$. u_t is the measurement error, independent from x_t and normally distributed with mean 0 and variance δ^2. Suppose there are two farmers *i* and *j*. We assume that farmer *i* can still observe the production gain perfectly but the effort is communicated with noise. The posterior beliefs of farmer *i* will be $f(B|(x_t + u_t))$. Since $f(x)$ is $N(B, \sigma^2)$ and $g(u)$ is $N(0, \delta^2)$ and *x* and *u* are independent $(\text{cov}(x,u) = 0)$ then $h(x,u)$ is $N(B, \sigma^2 + \delta^2)$. The farmer's beliefs about the benefits of the technique are updated in the following way (Berger, 1985; Udry and Conley, 2001):

$$\beta_t(x) = \frac{\sigma^2 + \delta^2}{\tau_{t-1}^2 + \sigma^2 + \delta^2}\beta_{t-1} + \frac{\tau_{t-1}^2}{\tau_{t-1}^2 + \sigma^2 + \delta^2}x_{jt} \tag{3}$$

where the farmer attaches more weight to his own prior beliefs.

The updating of beliefs may be dependent on the relationship with the experimenting farmer as suggested by Udry and Conley. In the next step we try to formalize this by making the measurement error dependent on the relationship between the farmers. We assume the variance of the noise is small if farmers have a close relationship, e.g. if they know each other since a long time and trust each other very well. The variance of the noise is large if the information is revealed by farmers with whom farmer *i* has no special relationship. As such the outcomes of farmers are weighed differently, depending on the value farmers attach to information from different sources. In the limit, δ^2 goes to zero for the closest farmers and it goes to infinity for farmers not known at all.

$$\text{For } \delta^2 \to 0: \ \beta_t(x) = \frac{\sigma^2}{\tau_{t-1}^2 + \sigma^2}\beta_{t-1} + \frac{\tau_{t-1}^2}{\tau_{t-1}^2 + \sigma^2}x_{jt} \tag{4}$$

$$\text{For } \delta^2 \to \infty: \ \beta_t(x) = \beta_{t-1} \tag{5}$$

[6] Udry and Conley (2001) find for pineapple growing farmers in Ghana that information flows via sparse networks rather than through the whole village.

When δ^2 goes to infinity, farmer *i* does not use the information revealed by the experimenting farmers and his prior beliefs are not updated.

2.2. Social learning test: problems and solutions

The theory suggests that farmers learn from each other, but not from all farmers in the same way. Information does not flow perfectly between farmers and in some networks the variance of the noise is smaller than in others. In our empirical work we aim to test whether social learning between farmers in different types of networks exists. In a first step we analyse behaviour with respect to technique adoption directly. Secondly, the existence of an information channel between farmers may give rise to social multiplier effects in banana output. If agricultural methods are correctly applied en well transmitted they may cause output to be related. We analyse whether group achievement actually affects individual achievement. If positive, it proves the existence of social effects amongst farmers which can be learning effects but might as well be the result of mimicking. To distinguish learning from mimicking we look into the impact of network members' information acquiring actions on a farmer's achievement (as advocated by Foster and Rosenzweig, 1995).

In most of the literature the farmers assumed to be group members are all farmers living in the same neighbourhood, the same village, district or belonging to the same ethnic group. However, these are all expected interaction groups. Because of data limitations many authors are forced to use expected interaction groups. Recently, some authors collected information on the exact interaction groups farmers belong to. Udry and Conley (2000) collected data on the information network of pineapple growing farmers in Ghana e.g. by asking for a list of persons whether and how many times they talked to a certain person. Bandiera and Rasul (2001) use a dataset on adoption of sunflower in Mozambique. They asked how many other adopters farmers knew e.g. within their family, neighbourhood or group of church members. We have information on actual social groups a farmer is part of, e.g. social support networks, kinship networks and geographical networks. We will show in part three that these groups are good proxies for a farmer's information network. Moreover, since we have data on all farmers in one village, we know the exact exogenous characteristics of all group members. Hence we can correct for exogenous social effects which could

induce group members' output to be related and lead us to falsely believe that social multiplier effects exist.

For those group members that are living in the same village as farmer *i*, we will assume they have the opportunity to observe closely the proceedings on the other group members' farms. Moreover, if good outcomes are observed, farmers inform each other on what was done to achieve this outcome, e.g. what kind of methods were used and how exactly they were applied. Hence we have found a good way to test learning from group members. If one of the network members *j* finds an effective technique and the exact way of applying it, which results in higher banana output, he can pass this information on to other network members *i*. If network member *i* rightly obtains and uses the information, we expect output to rise on network member *i*'s farm too. So the social learning test will take the following form:

$$y_{i,t} = f(Z_i, \beta_{i,t-1}) \tag{6}$$

with $y_{i,t}$ is the benefit of farmer *i* in period *t*. Z_i is a vector of the characteristics of farmer *i* and $\beta_{i,t-1}$ is farmer *i*'s belief about the benefit (or expected benefit) of applying a technique. It contains a history of private information the farmer received and information received from other farmers' experiences. As we have showed in (3) $\beta_{i,t-1}$ is a function of the beliefs of farmer *i* prior to experimentation and of the average benefit on the experimenting farmers' fields in period *t-1*. In the test we will assume that his prior belief about the benefit is zero and only include the last part, the average benefit on his group members' fields ($\bar{y}_j$).

$$y_{i,t} = f(Z_i, \bar{y}_{j,t-1}) \tag{7}$$

We assume the process takes the following chronology: in period *t-1* farmer *i* receives information about $y_{j,t-1}$ on his network member *j*'s farm (for all J network members) and asks what kind of techniques farmer *j* used in period *t-2* to obtain this result. If farmer *i*'s updated beliefs about the true benefit of the technique are larger than zero, he adopts the technique in period *t-1* and dependent on how well the information was transmitted, he will obtain increased output results in period *t*. So we assume $\bar{y}_{j,t-1}$ to have a positive effect on $y_{i,t}$ if farmers really learn from each other and to have no significant effect otherwise.

What we aim to test is an example of what is described in Manski (1993, 2000) as endogenous social effects. However, the identification of endogenous social effects may pose some serious problems. There have been many debates on how exactly society affects an individual and whether "real" social effects actually exist. Another problem was how to model them empirically (Brock and Durlauf, 2001). Manski describes and formalizes three hypotheses that exist to explain the observation that individuals in the same social group tend to behave similarly, only one of which are true social effects. Firstly, endogenous effects can play a role, where individual behaviour varies with or is affected by group behaviour. So there is an endogenous effect if, ceteris paribus, individual outcome tends to vary with group achievement. Secondly, there can be exogenous effects, where an individual's behaviour varies with exogenous characteristics of the group he is part of. For example, if outcome varies with the socio-economic characteristics of the group, there are exogenous effects. Finally, correlated effects can explain individual behaviour, where individuals in the same group behave similarly because they have similar individual characteristics or face the same environment. For example, if outcome varies similarly across group members because they face the same production constraints, there are correlated effects. All three effects have different policy implications. Manski gives the example of high school students. If a tutoring programme is implemented for some of the students, then there can be important social multipliers if there are endogenous effects. Then the programme does not only affect the achievement of the tutored students, but indirectly also affects the achievement of other students in the group. The other two effects do not generate this social multiplier effect.

Manski tackles the problem of the identification of the endogenous social effects, which is not straightforward. A serious problem encountered in trying to identify social interactions, is the so-called reflection problem. The reflection problem arises because the behaviour of the farmers in the reference group affects the behaviour of an individual farmer in that group but the behaviour of that farmer affects group behaviour. So there is a strong simultaneity problem. One of the possible solutions Manski offers to solve the simultaneity bias is to make the model dynamic and assume a lag in the transmission of social effects. So including lagged group behaviour instead of contemporaneous group behaviour can be necessary though not sufficient for identification. Only if the process of social effects is observed out of

equilibrium including lagged group behaviour may be a solution to the identification problem. Moreover, the timing of the lag has to be known. We claim that both conditions are present in our data set. Since not many farmers are applying the techniques yet and feel they are not yet necessary on their farm or too difficult, this is clear evidence of an out-of-equilibrium situation. We also claim the one year time lag is appropriate since the banana plants take approximately one year to become fully grown and flowering. Our results will even suffer less from the simultaneity bias since we have exact data on all group members and can exclude the observation of farmer *i* from the lagged group average.

So equation (7) represents the pure endogenous effects model where it is assumed that the coefficients of the exogenous effects are zero. In the presented form the test has not only been used to analyse social interactions in the adoption and effects of a new technology or in output performance, where the channel of social interactions is mainly information but in other types of behaviour as well, where the social interaction is exhibited via norms. For example, Borjas (1992) tested social interactions via the effect of average earnings within the same ethnic group as the parents, on current earnings of the children, via peer pressure of that group on the parents. Bertrand, e.a. (2000) used a similar basic test to analyse welfare use and how being part of a social group may inhibit mobility. Krishnan (2001) analysed the fertility behaviour of Indian women.

Obviously, the assumption of zero exogenous effects in equation (7) is a very strong one. Many authors are forced to make this assumption because data on exogenous characteristics of group members are not readily available. It is a typical omitted variables problem which results in biased estimates of the coefficient on $\bar{y}_{j,t-1}$. Therefore we drop the assumption of zero coefficients for the exogenous effects by including them explicitly in the test. The existence of correlated effects is partly taken into account through including the farmer's individual characteristics and other correlated effects are assumed zero since all surveyed farmers live in the same village and face the same environmental factors and constraints. It seems evident to include the factors that explain exogenous social effects since outcome may be driven by group members' characteristics and not by group behaviour as such. If we exclude the exogenous effects, the endogenous effect may be overestimated.

This argument holds even stronger in the case where social groups are endogenously formed, such as social insurance networks (De Weerdt, 2002). So far we assumed that the rules of group formation do not have any effect on the identification of the social effects. But when possible, individuals endogenously sort themselves into groups. For example, farmers will try to link up with farmers who have certain characteristics or abilities. The sorting of social insurance groups is a purely endogenous process. For our sample De Weerdt (2002) finds that characteristics such as kinship, clan, distance, education and wealth determine social insurance group formation. When endogenous matching takes place, there is potential for self-selection bias[7] and we may misinterpret the endogenous social effects. The variables that drive group formation may also drive farmers' outcomes but there are no endogenous social effects present. Therefore we will have to separate the exogenous group characteristics from the group behaviour itself and our test becomes:

$$y_{i,t} = f(Z_i, \overline{y}_{j,t-1}, \overline{Z}_j) \tag{8}$$

If the coefficient $\overline{y}_{j,t-1}$ is non-zero, then there are endogenous social effects. If the coefficient vector of $\overline{Z}_j$ which is a vector of average group characteristics proves significantly different from zero, there are exogenous social effects. A non-zero coefficient on Z_i indicates the existence of correlated effects.

Identification of both effects is strongly dependent on the correct identification of the reference group. Most studies use a variable that captures average behaviour of the whole expected reference group which is used for all observations in that group as an explanatory variable. Hence the average behaviour and average exogenous characteristics are not individual specific variables but lie at the village or neighbourhood level and are therefore only estimates of individual specific variables. We have information on individual specific reference groups and on all members' exogenous characteristics since all were included in the sample. Not many authors have the advantage of working with data on exact group composition (Udry and Conley, 2002; Bandiera and Rasul, 2001) and average group behaviour but to our knowledge there is none so far that uses information on correct individual specific exogenous group characteristics.

[7] Brock and Durlauf (2001) show that self-selection may actually facilitate identification. Self-selection

3. Information flows in Nyakatoke

Nyakatoke is a small village close to Bukoba town, in the Kagera region of Tanzania, west of Lake Victoria. The original tribe living in this area are the Haya. In this region the cooking banana is the main staple food. Average annual consumption per person is between 250 and 350 kg in the Lake Victoria region (Mbwana e.a., 1998). It is not only the most important staple food but also an important source of income earning for small subsistence farmers. However, for some years the productivity of the indigenous banana trees has been declining steadily mainly due to increased incidence of weevil pests and panama disease and the overall declining productivity of the home garden. The decline in productivity is due to a decline in soil fertility, which in turn is principally caused by increased population pressure. A second reason is the leaching of nutrients (or the natural soil depletion). The productivity problem is most severe in the high rainfall zone close to the lake (Baijukya and Folmer, 1999). There are strong differences between the banana yields in the different agro-economic zones of Kagera. In Eastern Kagera (where the survey village is located) the average yield is only 3100 kilograms per hectare, whereas in Central Kagera it is 6800 and in Western Kagera it is even 7500 (ARI Maruku, 1999). NGOs are active in trying to introduce new kinds of bananas that are more resistant to weevils and the panama disease. The new kinds of bananas are proven to give larger bunches and unlike with high yielding rice varieties that give higher output only when certain weather and input conditions are met, the ideal circumstances to grow new bananas are equal to those of the indigenous ones. No additional inputs or “good” weather are required. Even when farmers are aware of this, they are reluctant to adopt these new kinds. They cling to the old kinds because of their better taste. Hence there has been increased focus on the adoption of techniques that should prevent or mitigate the effects of weevil attacks and the panama disease and increase the productivity of the indigenous banana plant.

The data set contains information at the household as well as the individual level. Household heads were interviewed on household matters and all men and women in the household who earned their own income where interviewed separately by an enumerator of the same gender. At the household level we collected information on

---

may induce the sort of non-linearities that generate identification of the endogenous effects.

household characteristics such as land ownership, assets, quality of housing, food consumption, expenditure and household demographics. At the individual level the we have information on income, harvest, time allocation, social network members, kinship links, information flows, etc. Moreover, we have not taken a sample from the village but the survey includes all 118 households living in the village. This is advantageous if we want to analyse social effects. We know an individual's intra-village social links and we have full information on each of them. In the remainder of this section we will look at various aspects of information flows in the village.

Although the output decline of the indigenous banana is perceived as a serious problem[8] the reaction of farmers both in terms of adopting new types of bananas and in terms of using productivity enhancing techniques is still limited. In the villages located closer to the lake border in the high rainfall zone of Kagera, the soil is more depleted and most of the farmers have adopted new kinds of banana. But in our survey village the adoption of new types of bananas is marginal and mainly for experimental reasons[9]. Therefore we will focus on the process of information flows in general and with respect to technique adoption in particular. In the first paragraph we analyse knowledge and adoption of techniques. In the second paragraph we look at agricultural advice in general, namely where farmers get informed on agricultural practice. In the last paragraph we determine the characteristics of those farmers that pull information in the village. Starting from those farmers information can flow to other farmers in the village.

Our aim is to define a farmer's information network. However, we do not have sufficient direct data on it. Therefore we will look into the characteristics of the farmers one has an information link with. As such we are able to define other social networks that can be used as proxies for the information network. In order of exogeneity of the group, we will use a farmer's kinship network, his neighbours and his social support network. A household's kinship network is formed by all other households one of the household members has a blood bond with, up to the third

[8] In annex table A.1 we find that 26 percent of the farmers believe their output decreased slightly and 45 percent believes it decreased strongly. The productivity decline is seen as one of the major economic problems in the village.

[9] In annex A.2 we find that only 22 households have adopted other than the indigenous eastern African highland banana. Most of the households have planted them for experimental reasons. They usually have only one or two plants. In annex A.3 we find evidence that banana adoption is only done by the better educated and the wealthier households. So adopting new varieties is not (yet) a survival strategy.

degree. To form a network of neighbours we have taken 300 metres as the threshold for the members of two households to be neighbours. The average distance in the whole village is 523 metres. The last group is a household's social support network. It includes all the households one of the household members mentioned he could rely on in case of need and vice versa.

3.1. Adoption of productivity enhancing techniques

In this paragraph we discuss technique adoption and diffusion. Farmers were interviewed about their knowledge, adoption and information source of ten techniques. The techniques are explained in annex table A.4. Some of these techniques are very straightforward and easy to apply, others are more time consuming or need specific knowledge. Still others require manure or fertiliser which makes their adoption dependent on the cash situation of households. Hence technique adoption is a heterogeneous variable. It is not straightforward to construct one cumulative variable or an index that captures household adoption. Therefore we will look at the ten techniques separately.

In table 3.1 we show the percentages of household heads who know and use a technique. There is a large discrepancy between knowing and actually adopting the technique. The main reasons mentioned for not applying a certain technique were that farmers lack the money to buy the necessary inputs, that farmers do not believe in the effectiveness of the technique or that they have only recently learned it and did not start applying it yet. Some answered the techniques were too difficult (annex tables A.5). The capital constraint is only important for those techniques where manure or fertiliser is used. Not many farmers own cows or other animals, which makes manure a scarce and relatively expensive input. Also chemical fertiliser is both an expensive and not readily available input.

Some of the techniques were already known by the farmers when they started their farm because they saw their parents practising it. Other techniques were only recently introduced by extension agents. However, farmers do not only learn from extension agents, they also learn from other farmers outside or inside their community. In annex figure A.1 we represent all farmers who know at least one technique and where they acquired the knowledge. We can clearly observe two layers: included in the first layer are those farmers that pull information into the village. The upper layer represents

intra-village learning. It includes farmers who receive their information from other community members. The graph gives a good view of the flow of information entering and spreading through the village. We observe that a lot of farmers get their information from outside the village: from extension workers or from friends and relatives living in other villages[10] or even from farmers they do not have a connection with. Most farmers appear to be linked directly or through only one other farmer to an outside learning source.

Table 3.1: Knowledge and use of productivity enhancing techniques[a], n=113

Techniques	Knows the technique (%)	Uses the technique (%)
Special way of digging the hole	61	27
Applying fertiliser/manure	69	28
Hot water treatment of the stem before planting	18	3
Dipping stem in insecticide solution	10	0
Mulching 1 meter from stem	10	1
Trench-manure (water conservation)	24	7
Paring	48	39
Desuckering (3 plants per stool)	35	19
Harvest hygiene	35	12
Weevil trapping	54	22

[a] for a more detailed explanation of the techniques, see annex A.4

In table 3.2 we show learning sources and the percentage of the total number of techniques known that is learned from that type of source. We observe that 57 percent of the techniques known have been pulled into the village and 43 percent have been learned from another farmer or group of farmers in the same village. We notice that a higher percentage of the techniques that farmers know is actually applied when farmers have learned it from an outside source. Clearly, this is a self-selection problem. Farmers who learn techniques from outside the village have chosen to put some effort into looking for information and to undertake transaction costs to learn the technique with the specific aim of applying it. Whereas farmers who learned techniques from other farmers in the village, may have accidentally learned it while visiting each other or discussing agricultural issues but do not believe it necessary to adopt.

[10] In the case of whom techniques are learned from, the "outside relatives" option sometimes includes relatives who are already dead such as parents or grandparents.

Table 3.2: Learning and diffusion of techniques, n=397[a]

Source	% of techniques learned from source	% of known techniques used
Information enters village via:	*57*	
Farmers outside community	6	55
Formal extension (NGO/government/seminars)	48	47
Other (Self-taught, school)	3	50
Diffused further via:	*43*	
Farmers inside community	38	43
No specifically identifiable person[b]	6	32
All techniques learned	*100*	44

[a] Several observations per household possible, depending on the number of techniques known by the respondent (household head); [b] If the teacher of a technique could not be specified, he was either dead or it concerned a group of farmers.

With respect to the intra-village learning links, we are interested in whether they are in any way determined by the relationship between both farmers. We test whether learning links exist between random villagers or whether they exist significantly more within certain groups. In table 3.3 we present some information on learning clusters. It shows that the households one learns from are significantly more of the same social group than those one does not have a learning link with. For example, 45 percent of the farmers one has learned techniques from belong to kin-related households. Of all the households one does not learn from, only five percent are kin-related (which is the average percentage of kin-related households in the village). Consequently, we assume that these social groups, on each of which we have full information for all respondents, can serve as proxies for farmers' actual learning networks (for which we do not have data for all respondents, since the question on techniques learned was only answered by household heads who knew that particular technique).

Table 3.3: Learning links within social groups

Social groups based on:	Percentage of households of the same social group if		
	Learning link	No learning link	Significance
Kinship	45	5	***
Clan	38	9	***
Religion	60	35	***
Distance (metres)	249	522	***
Insurance group	55	7	***
Observations (nr)[a]	40	14002	

[a] All households (119) linked with all other households results in 14042 observations: we find a one directional learning link for 40 combinations; *** significant at 1%

The descriptive statistics in table 3.3 are confirmed by the regressions in annex table A.6, columns (1) and (2). The left hand side variable is one if two households have a learning link and zero otherwise. We used relational variables to explain the link, such as whether both household heads received lower primary education or whether a

kinship or a social support link exists. Apparently, whether the household heads are of the same gender weakly affects the possibility that a learning link exists, but whether both households are kin-related, belong to the same social insurance network or live closer together are strongly significant in explaining the existence of a learning link.

3.2. Advice seeking behaviour

In another part of the survey we did not restrict information flows to knowing or adopting specific techniques but we inquired very openly about farmers' agricultural advice sources[11]. The question was framed as "whom would you go to for advice if there was a problem with your crops?". In annex figure A.2 we show all farmers in the village connected to their hypothetical adviser[12]. The graph differs substantially from the graph in annex figure A.1 where it concerns the number of layers. With respect to hypothetical advice seeking, there are more layers of intra-village links: a number of farmers are many steps away from an outside advice source.

Table 3.4: Hypothetical information links and social groups

Social groups based on:	Percentage of households of the same social group if		
	Advice link	No advice link	Significance
Kinship	38	5	***
Clan	26	9	***
Religion	50	35	***
Distance (metres)	295	523	***
Insurance group	76	7	***
Observations (nr)	74	13968	

*** significant at 1%

With respect to social groups the hypothetical advisers belong to (table 3.4), the same picture emerges as what we found for actual learning links. In annex table A.6, columns (3) and (4) we find what type of relationship determines whether a farmer mentions another farmer in the village as his or her hypothetical advice source. The group dummies (farmers belong to kinship related households, distance between households, farmers belong to households in the same social support network) significantly affect the probability of a hypothetical advice link between two respondents. Being of the same gender has a strong positive effect on the probability

[11] Unlike the question related to the knowledge and adoption of techniques, the question on advice sources was asked to all individuals in the sample (household heads and spouses/children who earn an own - agricultural - income).

[12] These are the "opinion leaders" in Rogers (1995). Opinion leaders can be identified by the sociometric method, which entails exactly the same question as was used in our survey (whom would you –actually or hypothetically- go to for advice or information).

of advice links. The gender effect indicates that women consult other women and men consult men. If both farmers received lower primary education the probability that they mention each other as an adviser is significantly lower than if only one of both would have reached this level of education or none of them did. Further, table A.6 shows that belonging to the same household increases the probability of a hypothetical advice link. Thus farmers will consult other farmers in the household. This is suggestive of the existence of intra-household information flows[13].

From 3.1 and 3.2 we conclude that information will not be spread to all farmers in the village but it flows via networks[14]. The tables support the idea that social groups based on kinship, neighbourhoods and social support systems are good proxies for a farmer's intra-village agricultural information network. Additional rejection of the assumption that the whole village constitutes one information network lies in the following observation. We listed ten persons and asked for all of them whether the respondent had gone for advice to this particular person the year before the survey. Persons included were for example the local extension officer, the innovator of the village, some social support network members and one randomly drawn male and female farmer. The percentage of advice farmers obtained from a randomly drawn person from the village was close to zero[15].

3.3. Information pullers

From the tables we can conclude that there is a lot of information being transferred between the farmers of a social group. However, the link is often one-directional. There are some key group members who are asked for advice but who get their advice from outside the group or the village. Table A.7 in annex shows some regressions to identify the characteristics of farmers that have an outside information link (i.e. farmers who pull information in the village). With respect to learning techniques (which is regressed at the household level) we find that education has a weakly

[13] Mostly the direction is from women asking their husbands for advice and not the other way around. We use this evidence later to justify the use of the network of the whole household rather than restrict the information network to the individual farmer's own network.

[14] "*Individuals tend to be linked [in information networks] with others who are close to them in spatial and social distance*", (Rogers, 1995,p.311) which is very close to the results found in De Weerdt, 2002. He finds that physical distance is very important in determining whether two individuals have a social insurance link. Moreover religion and education are significantly more the same in a social insurance network compared to random clusters drawn from the village.

positive effect on having learned techniques from a source outside of the village. A household head who completed primary school has a 19 percent higher probability of having learned techniques from an extension officer. But the most significant factor is the household's landholdings: each hectare more results in a 15 percent higher probability of learning from an extension officer. At the individual level (having outside advice sources) the regressions mainly reveal that gender and secondary education strongly increase the probability that the farmer mentions an outside advice source. Men have a nearly 20 percent higher probability of contacting an outside advice source or an extension officer. Some primary education has a slightly positive effect but some secondary education increases the probability of seeking advice from an outside source or more particular from an extension officer by 35 and 42 percent respectively.

We can draw two main conclusions from the preceding analysis. Firstly, information flows appear to be tied to social group. Social linkages determine strongly the probability that an agricultural information link exists. Secondly, those that initially pull information in the village are more often higher educated men living in households with large landholdings. Hence there may be a pool of "teachers" within each social group and the social effects may be different with respect to position in the group. Farmers who have outside information sources might not learn from other group members or the highest productive farmers in a group may prefer to seek information outside their group. We will take these observations into account in the next section.

In section four we test how adoption behaviour in a farmer's social group affects his own behaviour. The final question is whether information flows actually induce social multiplier effects in banana cultivation. We analyse whether group members' outcomes are related taking into account that there may be a direction to the flow of information in the group.

[15] The sample probability that there is a hypothetical (ex ante) advice link between two villagers is close to zero percent (0.159%). The percentage of advice asked to a randomly drawn female farmer was 0 percent, to a male farmer 4 percent.

4. Empirical results

The empirical section consists of two parts. First we will test whether behaviour with respect to technique adoption in a farmer's social group actually affects his own behaviour. Secondly we test whether information flows can be the underlying mechanism for endogenous social effects by analysing the relationship between a farmer's banana output and average output performance in the social group he is part of.

In section two we showed how farmers update their beliefs about the expected benefits of applying a certain technique, based on the experience of other farmers. The updating mechanism is dependent on the relationship between farmers. A farmer attaches more weight to information received from "group members" than information received from a random farmer in the village. In section three we found for the survey village that information flows via certain linkages indeed and does not reach all farmers of the village in the same way. Farmers who are part of the same social group not only visit and talk to each other more frequently[16], thereby providing opportunities to observe more closely the proceedings on the other's field, but the variance of the noise in transmitting information is smaller. Hence they draw more exact conclusions on the benefits of agricultural practices. As such, improved output results on the farm of one of the members of farmer *i*'s social group can have positive effects on the outcome of farmer *i*.

From section three we concluded that groups based on kinship, distance and social support systems are good proxies for a farmer's information network. Hence, these groups will be used throughout the analysis. Unlike the kinship and distance based groups, the social support group is endogenously formed. Farmers choose whom to link up with in this type of group. Hence we have to be careful in interpreting the social effects based on social support groups. It may well be that the variables that drive group formation also drive farmers' behaviour. In that case there may be exogenous social effects but no social multiplying effect (endogenous social effects). The problem is less stringent when we correct for exogenous group characteristics. Since we have full data on all the members of a farmer's kinship, neighbour and

[16] Sometimes a farmer's social network does include persons who live further away and who hardly ever meet each other but for the analysis we only included those network members living in the same village.

social support group, who are living inside the village, we can correct for exogenous group characteristics and safely test the hypothesis of endogenous social effects in growing bananas. Before we continue we present some information on the banana growers in the village. Table 4.1 shows some averages.

Table 4.1: Some characteristics of banana growers

Variables	Obs.	Average
Individual characteristics:		
Male growers (%)	117	50
Age	117	43
Some primary education (% reached standard 1 to 4)	117	77
Completed primary education (% standard 5 to 7)	117	56
Number of banana sellers	117	95
Output value of banana sellers (in Tsh[a])	95	6841
Of all households with at least one banana grower:		
Kibanja (=most fertile plot, in hectare)	101	0.5
Total land holdings (hectare)	101	1.2
Adults present in the household (> 15 years)	101	2
Social groups of banana sellers (nr of persons in group):		
Number of kinship members	88	8
Banana selling kinship members 1999	88	5
Number of neighbours (living at less than 300 metres)	95	32
Banana selling neighbours 1999	95	21
Number of social support network members	95	10
Banana selling social support network members 1999	95	7

[a] Tanzanian Shilling (1 US$=+/-800Tsh)

The gender ratio is very equal: 50 percent of all banana growers are male. Although bananas are grown by all households, only 95 persons (81% of all banana growers) mentioned selling bananas as an income earning activity. For banana selling respondents we have information on total output values, both what was sold and what was consumed by the household. Due to the design of the survey we do not have information on output values for those households that did not mention banana cultivation as a source of cash income[17]. Household land is on average 1,2 hectares, nearly half of which is cultivated with bananas (kibanja). With respect to social groups, there are some households that do not have any kin related households in the village. For two households to be neighbours we used a threshold of living within 300 metres from each other. Obviously 300 metres is an arbitrary measure. We chose it because it is below the average distance in the village (523 metres) and if we were to

[17] The survey aimed to capture only the income earning activities individuals where engaged in and no information was collected on output values of crops grown only for household consumption. For households that did not sell bananas output values are missing.

opt for a lower threshold some remote households would be dropped from the analysis.

4.1. Direct test of social effects in technique adoption

Before we test whether social interactions in banana growing exist we test the underlying assumption of information flows. We analyse whether adoption of a certain technique is dependent on the number of other adopting farmers in a group. In table 4.2 we show the regression results explaining adoption of techniques by household characteristics, exogenous group characteristics and the number of adopting group members.

Table 4.2: Effect of group use of a technique on household adoption, pooled techniques (marginal effects)

Dependent variable: use of technique (0/1)	**Kinship**		**Neighbours <300m**		**Social support network**	
	Coeff.	Standard error[a]	Coeff.	Standard error[a]	Coeff.	Standard error[a]
Household characteristics						
Total land (hectares)	0.039	0.008***	0.033	0.009***	0.023	0.009***
Adults present	-0.012	0.010	0.005	0.010	0.014	0.011
Sex of household head	0.000	0.029	0.062	0.021***	0.062	0.022***
Head some lower primary	0.039	0.034	-0.095	0.050*	-0.133	0.055**
Head completed primary	0.075	0.030***	0.088	0.027***	0.092	0.026***
Age of household head	0.000	0.001	-0.001	0.001	-0.000	0.001
Group characteristics						
Nr of group members using technique	0.054	0.007***	0.018	0.002***	0.047	0.006***
Nr. lower primary	0.010	0.013	-0.019	0.009**	0.012	0.011
Nr. higher primary	-0.009	0.011	0.007	0.005	0.024	0.010**
Nr. male headed	-0.027	0.011**	0.009	0.008	-0.040	0.009***
Average age	-0.005	0.002***	-0.009	0.004**	-0.001	0.002
Constant	-0.172	0.963	2.189	1.991	-2.543	0.932***
Observations	849		929		929	
Pseudo R^2	0.186		0.192		0.199	

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

[a] Technique adoption was asked for a series of ten techniques; standard errors are adjusted for household clustering

The dependent variable is one if the household uses the technique. Since we asked about ten different techniques which are difficult to combine in one variable, there are ten observations per household. The standard errors presented in the table are adjusted for household clustering. The household level explanatory variables include *total land holdings* measured in hectares, number of *adults present*, *sex* of the household head, his *age* and two dummies to capture educational levels. One dummy captures *some primary schooling* (first four years) and the other captures *higher primary schooling*

(years five to seven completed). With respect to the social groups the household is part of, we include the *number of group members adopting* respective techniques. Other than technique use in the group we included the *number of household heads with some primary schooling*, the number of heads *with completed primary schooling*, the group *members' average age* and the *number of male headed* households in the group. We repeat the test for the three different social groups.

At the household level we find that the size of the farm is a significantly positive determinant of technique adoption. Individual education effects consistently exist only for farmers who completed primary school. We do not find consistent effects with respect to lower primary education and the sex of the household head. From the neighbours and social support network regressions we observe that male headed households have a significantly higher probability of technique adoption. This is confirmed by the results in annex (table A.8) where adoption is shown without taking into account social effects. However, because social effects are not taken into account the individual effects are slightly overestimated.

Although the coefficients differ we find the effect of the number of techniques used by group members to be positive and significant for all three social groups. One adopting kin member or social support network member more increases the probability of a farmer adopting a technique by 5 percent. Neighbours who have adopted a technique show a lower impact of 2 percent. Unfortunately we do not have information on the year of adoption which forces us to analyse contemporaneous behaviour. Hence the results may suffer from a simultaneity bias. The exogenous group effects strongly differ between groups. Now that social effects in technique adoption have been proven to exist in our sample, we continue with the actual social multiplier test.

4.2. Indirect test: endogenous social interactions in banana output

First, we present the form of the test and describe the variables used. For the test we will first assume exogenous social effects to be non-existent and afterwards allowing individual output to vary with both group behaviour and exogenous group characteristics. It will clearly show the limitations to the conclusions we can draw from the first test. With respect to the observations included in the analysis, we will take into account the findings from section three. It became clear that there are

differences between the farmers regarding their information behaviour. While some pull information inside the village, others tap their information from other farmers inside the village. This behaviour may affect the direction of the externalities on individual outcome. Therefore we will correct for the relative yield position of a farmer within his social group. We assume that farmers prefer to learn rather from their high yield group members and the latter will learn from outside sources. Including the "teachers" may underestimate the social effects. Therefore we exclude the two most productive (measured by the output per hectare) group members from the analysis.

To test social interactions in banana output we use the following log-linearized Cobb-Douglas production function:

$$\log(Y_{i,t}) = \alpha + \beta \log(Z_i) + \chi \log(X_i) + \delta \log(\overline{Y}_{-i,t-1}) + \varepsilon \log(\overline{X}_{-i}) + u_t \qquad (4.1)$$

$Log(Y_{i,t})$ is the logarithm of the output value of the banana harvest of farmer *i* at time *t*. Z_i is a vector with characteristics of the household farmer *i* belongs to. Conventional inputs in the production process are area and labour allocated to the cultivation of banana trees. The number of hectares cultivated with banana trees is usually restricted to the plot around the house[18] (*kibanja*), which is the most fertile of all plots. Other than banana trees, the kibanja is also cultivated with coffee, maize and beans, all intercropped. We take the area of the kibanja as exogenous. There is a very strong case to do so because the area where the survey was done is characterised by a high population density. The house with the kibanja area around it is mostly constructed at marriage and the land is part of the husbands' family land. There is not much opportunity to expand the kibanja area since neighbours' kibanjas are next to it[19]. Some households do grow banana trees on less fertile grounds but mostly the banana area is restricted to the kibanja area, which is in principle not expandable. Moreover, the banana plant is a quasi-permanent crop. It takes up to one year for the first flowering of the mother plant. The mother spontaneously produces a daughter and a granddaughter and then dies. The banana field is easily maintained and the

[18] Called the kibanja in Swahili, see Mitti and Rweyemamu (2001), p.15; Baijukya and Folmer (1999), p.45 for characteristics of the kibanja and other plot types and the type of crops grown on different plot types.

[19] A special type of plants marks the borders of the kibanjas, otherwise it would look like one very large banana field.

farmer even has to remove excess suckers. Hence the kibanja area is not a choice variable. We do not have individual information on the hectares cultivated with bananas. This information would be nearly impossible to collect since husband and wife often care for all the banana trees in the kibanja at the same time[20]. Men and women's banana trees are scattered all over the plot and often it is only the marketing of the crops that is done separately. Men are often responsible for the marketing of the beer bananas and women for marketing of the cooking bananas. But nothing prevents men to sell cooking or women to sell beer bananas. To correct for a possible gender bias we include the *sex* of the banana grower. Another household characteristic we will use as a regressor is the number of *adults* present to capture labour availability. X_i is a vector including individual characteristics of farmer *i*. Besides gender and age we include both education level dummies introduced in the preceding section: the farmer completed only lower primary school[21] or completed seven years of primary school. The logarithm is taken of all the continuous variables in the regression and dummies remain in the regression as levels.

$Log(\overline{Y}_{-i,t-1})$ is the logarithm of the *average banana yield* (output per hectare) at time *t-1*, of the households that are part of farmer *i*'s social group, excluding the output of farmer *i* himself. We use the average yield of the households that are linked to the household farmer *i* is part of. Therefore we have to make the assumption of intra-household information pooling: we assume that, if any household member receives information on banana cultivation, he will share this information with his fellow banana growing household members. Our data allow us to solve the reflection problem described in Manski (1999) by including lagged group yields and by excluding the individual from the average group behaviour. This is an important advantage of the data, which were collected with the specific aim to identify social groups within the village, a feature which most other datasets lack. Moreover, the data allow us to correct for exogenous group characteristics.

Any social effect can be (partly) caused by exogenous characteristics of the whole group that drive output. One should be especially suspicious of this phenomenon in

[20] When the men are mulching the kibanja they will mulch the whole plot and not just their own banana trees. When women are weeding their beans, which grow under the banana trees, they are in fact also weeding the banana field.

[21] Lower primary school is from first to fourth grade of primary school. Higher primary is from grade five to seven.

the case of endogenously formed groups such as the social support network. $\overline{X}_{-i}$ is a vector of group characteristics, such as average age of the group members and the stock of education in the group: the number of persons having some primary education and the number who completed primary education. If $\varepsilon \neq 0$, exogenous social effects exist. If $\delta \neq 0$ there are true endogenous social effects.

Table 4.3 shows the results of the social interactions test. We use a two-step Heckman selection model to correct for the sample selection problem (missing output data for non-selling banana growers). As first-step selection variables we use the number of children present in the household and the value of the household durables, which might influence the decision to sell bananas but not the output obtained by the banana grower. Table 4.3 is organised as follows. The first part shows the Cobb-Douglas estimation results for the case where exogenous social effects are assumed to be non-existent. The next part allows for both endogenous and exogenous social effects.

We exclude the two farmers with the highest yields relative in each group because they will be teachers or advisers rather than learners. This assumption is confirmed by our data. In annex table A.9 we find that the quartile of the highest performance farmers has usually a higher percentage of agricultural advisers, technique teachers and farmers with outside sources of information than the lowest performing quartile, especially in the exogenous groups kinship members and neighbours. Including them in the regressions can underestimate social effects. Taking into account relative yield positions is only possible due to the nature of our data since we know the exact group composition and have information on yields of all banana selling members. The cases where all farmers are included and where only the farmers in the highest yield position are excluded are presented in annex tables A.10a and A.10b.

From table 4.3 we find that the individual effects of land and lower primary education are fairly consistent throughout all the regression specifications. The social effects however, are quite different depending on both the group and the model we analyse. If exogenous effects are assumed zero we do not find any social effects in neighbourhoods but we do find significant endogenous social effects for kinship and social support groups. The table presents evidence that social capital truly is capital in the sense that it contributes to production (Narayan and Pritchett, 1999).

Table 4.3: Heckman two-step estimates[a] for endogenous and exogenous social interaction models, two highest yield farmers excluded

	Exogenous group effects assumed zero						Exogenous group effects nonzero					
Dependent variable:	(1) Kinship		(2) Neighbours <300m		(3) Social support network		(1) Kinship		(2) Neighbours <300m		(3) Social support network	
Log(banana output value)	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Grower characteristics												
Log(Kibanja)	0.621	0.302**	0.681	0.245***	0.618	0.252***	0.563	0.268**	0.778	0.215***	0.607	0.255**
Log(Adults)	1.279	0.662**	0.532	0.395	0.977	0.539*	1.013	0.582*	0.338	0.334	0.325	0.419
Male grower	0.162	0.506	-0.300	0.298	-0.092	0.348	0.205	0.428	-0.241	0.267	-0.122	0.307
Some primary school	0.913	0.622x	0.914	0.462**	1.085	0.510**	0.999	0.567*	1.226	0.431***	0.698	0.461x
Completed primary	0.333	0.556	-0.264	0.382	-0.145	0.420	0.005	0.556	-0.366	0.341	-0.138	0.374
Log(Age)	0.998	0.811	0.087	0.487	0.426	0.595	0.659	0.715	0.218	0.480	0.188	0.474
Group characteristics												
Log(Yield t-1)	0.603	0.322*	-0.380	0.542	0.456	0.248*	0.540	0.280**	-0.562	0.510	0.144	0.222
Nr. lower primary							0.498	0.303*	0.203	0.133x	0.329	0.152**
Nr. higher primary							-0.307	0.210x	-0.055	0.071	-0.177	0.143
Nr. male headed							-0.236	0.274	-0.097	0.110	-0.108	0.137
Average age							-0.19	0.033	-0.051	0.058	0.006	0.024
Constant	-3.328	6.019	10.788	4.879**	0.828	4.355	-0.630	5.530	12.417	5.706**	4.404	3.164
Mills lambda	1.421	1.197	-0.307	0.901	1.129	1.115	0.991	1.166	-0.120	0.855	0.543	0.737
Observations	79		107		97		79		107			97
Uncensored obs.	58		88		76		58		88			76
Wald chi²	28.37		29.32		29.78		43.94		53.86			49.10

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

[a] Selection effects not shown in the table

Including the groups' exogenous characteristics significantly changes the results. The single group for which the true endogenous social effects exist is that based on kinship links. For all three groups the stock of education in the group positively influences individual members' output but the coefficients are more significant and larger for kinship and social support groups than they are for neighbours. Besides private effects, there are public effects of primary education. This effect is comparable to the positive intra-household externality of a literate household member described in Basu and Foster (1998). In the social support group we observe that the private effect of primary education strongly decreases when corrected for education in the group. In the first set of models the individual education dummy captures the group effect. Apparently it is a characteristic that drives group formation and output at the same time: once corrected for education in the group the individual education effect becomes smaller and the endogenous social effect disappears. When we do not correct for exogenous characteristics the endogenous effect in social support groups appears misleadingly large[22].

Tables A.10a and A.10b in annex provide evidence to support the hypothesis that there is a direction to the social effects going from the better to the less performing group members. We clearly see that both the significance and the magnitude of the endogenous social effect increase when better performing farmers are excluded from the analysis.

A question that remains is why the social multiplier effect exclusively exists in kinship related groups. Possibly the answer lies in the fact that, before one can gain from another farmer's knowledge, information about agricultural practice has to be passed on very meticulously (and the variance of the noise in transmitting information

[22] As a robustness check we ran a fixed effects estimation by using the panel nature of our data. In 2000 harvest was reported three times, once for the first six months, then every three months and retrospectively for 1999. In the fixed effects regressions we included the number of adults present in the household and the average group productivity at *t-1*. These are the only variables in our sample that change over time though we have to add that the number of households where adults present changed was only marginal. We found positive but insignificant coefficients for lagged average group yield in all three groups. This may be due to the fact that the lags are rather short in order to measure any improved output effects in growing bananas. What we measure might simply be an approximation of the measurement error and not the change in output. Brock and Durlauf (2001) suggest that enough variation in the change in average behaviour is useful for identification. In slowly moving environments, the coefficient of the change in average behaviour may be difficult to estimate precisely.

is small). Merely observing what happens on another farm is not sufficient[23]; farmers need to know the exact way of how the good result was obtained[24]. Presumably kinship related farmers, e.g. parents and siblings take more time to explain the technology, than would neighbours or social support network members do. Moreover, between kin-related farmers there are fewer unobserved characteristics which can cause a higher variance of the noise in transmitting information. We explore the possibility of better information transmission further by analysing individual banana output with respect to information that was obtained by group members in period *t-1*. We replace lagged average group yield by the number of times group members had contact with an extension officer in *t-1*. We see that group members' contact positively affects individual output in kinship related groups which is suggestive of a better information transmitting mechanism in this type of groups. Moreover, the effects of technique use by group members (as shown in table 4.2) also proved the effects for kinship groups to be the largest. Hence we can conclude that in our sample information flows constitute the underlying mechanisms for the social multiplier effect in banana growing only for kin related villagers.

To summarize, we found that there are endogenous social effects only between kinship related farmers but they work from the better to the less performing group members. Additional to the endogenous effect there are public effects of education within all three groups although they are only weakly significant for neighbours. The finding that endogenous social effects that produce the social multiplier effect only exist within the kinship related groups in our sample may be due to better transmission of information in those groups. Supportive to this hypothesis we do find that information spill-overs (adoption of techniques, individual outcome effects of group advice seeking behaviour) are most prominent in kinship related groups.

[23] Munshi (2004) shows that information flows are weaker in a heterogeneous population when performance of a new technology is sensitive to unobserved individual characteristics.

[24] For example, a technique exists to prevent weevils from attacking the banana plant (Annex A.4). The farmer has to put some freshly cut pieces of the banana stem around the plant. Those pieces will attract the weevils before they reach the plant. The farmer has to clean these cut pieces from weevils attached to the bottom at regular times, e.g. at least once a day. But some periods of the day are better suited to clean these weevil traps, e.g. in the morning.

5. Conclusion

In Kagera, west of Lake Victoria, the banana is the main staple food and an important cash crop. Since many years bunches are decreasing in size due to weevil attacks and diseases. Farmer extension centres and NGOs are trying to introduce new types of bananas and productivity enhancing techniques to mitigate the negative effects. What these organisations often observe is that not many farmers appear to be interested in learning new techniques and even less so in actually applying them. However, the number of farmers that adopt techniques is not restricted to those that come to the centres. In our work we aimed to provide evidence that there are large externalities of the few farmers that learn agricultural methods directly from a farmer extension officer.

We have used a Bayesian framework to show the mechanism of updating of beliefs about the benefits of a certain technique based on the experiences by other farmers. The farmer updates his beliefs each time he observes the true benefits of the technique revealed by another farmer actually applying it. However, the farmer does not learn from all farmers in the same way but he attaches more value to the outcomes of farmers where information transmission is characterised by a lower variance of the noise.

This lead to a social interactions test of the Manski type: individual achievement varies with average group achievement (endogenous social effects) or with group characteristics (exogenous social effects). We have tackled most of the problems attached to this type of social interactions test. Firstly, we used lagged group achievement instead of contemporaneous achievement to solve the simultaneity problem. Moreover, we can claim that the process is observed out of equilibrium since many farmers do not apply the techniques yet. And most importantly, we know the exact interaction group composition and we have data on the achievement and the characteristics of all group members. Unlike many authors who are obliged to use datasets not designed to analyse social effects, ours was collected with the specific aim of analysing social groups. So we used individual specific groups and group characteristics for our analysis.

We used three different social groups. In order of exogeneity these are kinship related groups, neighbours and social support network members. Firstly, we found that

technique adoption is strongly affected by the number of adopting group members in all groups. But mere technique adoption does not prove to give rise to endogenous social effects in banana output. We found that endogenous social effects exist only within kinship groups and exogenous social effects of some years of primary education exist for all three groups, though the effects for neighbours are much lower and less significant than for the other types of groups. Following our theory we claim that the endogenous effects within families are (at least partly) caused by better-quality information transmission. To explore this route further, we analysed the effect of total extension contact by group members on individual banana output. We found that information spill-overs, in the sense that information obtained by group members actually affects other members' output, were only obtained amongst kinship related farmers.

Thus we can conclude for our survey village that information transmission only has social multiplier effects on farmer's banana output if it is passed on very meticulously. For positive group output effects it seems better to target farmers belonging to different kinship groups. Although the latter variable is a strong determinant of social support group formation and kinship related farmers tend to live in the same geographical area, it is not through living close together or through forming groups that the endogenous social effects in growing bananas came about but through long-term (family) relationships.

References

Agricultural Research Institute, (1999), "*Profitability of Crops in the Kagera Region*", in "Planning the Future: Past, Present and Future Perspectives of Land Use in the Kagera Region", ARI Maruku, Bukoba and Royal Tropical Institute, Amsterdam

Baijukya, F.P. and E.C.R. Folmer, (1999), "*Land Use in the Kagera Region: Definitions and Ratios*", in Planning the Future: Past, Present and Future Perspectives of Land Use in the Kagera Region, MARUKU Agricultural Research Institute

Bandiera, O. and I. Rasul, (2001), "*Social Networks and Technology Adoption in Northern Mozambique*", paper presented at NEUDC Conference, Boston University

Bardhan, P. and C. Udry, (1999), "*Development Microeconomics*", Oxford University Press

Berger, J.O., (1985), "*Statistical Decision Theory and Bayesian Analysis*", Springer-Verlag New York Inc

Bertrand, M., E.F.P. Lutmer and S. Mullainathan, (2000), "*Network Effects and Welfare Cultures*", The Quarterly Journal of Economics, vol.115, nr.3, pp.1019-1055

Besley, T. and A. Case, (1993), "*Modelling Technology Adoption in Developing Countries*", American Economic Review, vol.83, nr.2, pp.396-402

Besley, T. and A. Case, (1994), "*Diffusion as a Learning Process: Evidence from HYV Cotton*", Research Program in Development Studies (RPDS) discussion paper, nr.174

Borjas, G.J., (1992), "*Ethnic Capital and Intergenerational Mobility*", The Quarterly Journal of Economics, vol.107, nr.1, pp.123-150

Borjas, G.J., (1995), "*Ethnicity, Neighborhoods, and Human Capital Externalities*", American Economic Review, vol.85, nr.3, pp.365-390

Brock, W.A. and S.N. Durlauf, (2001), "*Interactions-based models*", in Handbook of Econometrics vol.5, Heckman, J. and E. Leamer (editors), Elsevier, Amsterdam

Case, A., (1992), "*Neighborhood Influence and Technological Change*", Regional Science and Urban Economics, vol.22, nr.3, pp.491-508

De Weerdt, J., (2002), "*Social Networks, Transfers and Insurance in Developing Countries*", PhD. Katholieke Universiteit Leuven

Ellison, G. And D. Fudenberg, (1993), "*Rules of Thumb for Social Learning*", Journal of Political Economy, vol.101, nr.4, pp.612-643

Ellison, G. And D. Fudenberg, (1995), "*Word-of-Mouth Communication and Social Learning*", The Quarterly Journal of Economics, vol.110, nr.1, pp.93-125

Feder, G., R.E. Just and D. Zilberman, (1985), "*Adoption of Agricultural Innovations in Developing Countries: A Survey*", Economic Development and Cultural Change, vol.33, nr.2, pp.255-298

Foster, J.E. and K. Basu, (1998), "*On Measuring Literacy*", Economic Journal, vol.108, pp.1733-1749

Foster, A.D. and M.R. Rosenzweig, (1995), "*Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture*", Journal of Political Economy, vol.103, nr.6, pp.1176-1209

Gale, D., (1996), "*What Have we Learned from Social Learning?*", European Economic Review, vol.40, pp.617-628

Krishnan, P., (2001), "*Cultural Norms, Social Interactions and the Fertility Transition in India*", paper presented at NEUDC Conference 2001, Boston University

Manski, C., (1993), "*Identification of Endogenous Social Effects: The Reflection Problem*", Review of Economic Studies, vol.60, pp.531-542

Manski, C., (2000), "*Economic Analysis of Social Interactions*", Journal of Economic Perspectives, vol.14, nr.3, pp.115-136

Mbwana, A.S.S., L. Ngode, K.V. Seshu Reddy and R.A. Sikora, (1998), "*A Guide to Growing Bananas in the Eastern African Highlands*", International Centre of Insect Physiology and Ecology

Mitti, R.D. and T.G. Rweyemamu, (2001), "*A Socio-Economic Profile of Nyakatoke*", Social Security, Poverty and Economic Reforms in Kagera Region, Report Nr. 1

Munshi, K., (2004), "*Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution*", Journal of Development Economics, vol.73, pp.185-213

Narayan, D. and L. Pritchett, (1999), "*Cents and Sociability:Household Income and Social Capital in Rural Tanzania*", Economic Development and Cultural Change, vol.47, nr.4, pp.871-897

Rogers, E.M., (1995), "*Diffusion of Innovations*", The Free Press, New York

Rweyemamu, T.D., (2001), "*A Socio-Economic Profile of Kaigi (Buhembe)*", Social Security, Poverty and Economic Reforms in Kagera Region, Report Nr. 4

Udry, C. and T. Conley, (2000), "*Learning About a New Technology: Pineapple in* Ghana", Yale Economic Growth Center Discussion Paper: 817

Udry, C. and T. Conley, (2001), "Social *Learning through Networks: The Adoption of New Agricultural Technologies in Ghana*", American Journal of Agricultural Economics, vol.83, no.3, pp.668-673

Annexes

Table A.1: Respondents observation of banana output, 2000 versus 1990

Output direction	Percentage
Increased strongly	10
Increased slightly	13
Not changed	6
Decreased slightly	26
Decreased strongly	45
Nr. of respondents	78*

*only 78 households of the 118 mention they grow banana commercially in 2000 and were already growing bananas ten years ago. The reasons may be that the household was not formed yet in 1990 or growing banana was not a commercial activity of any of the household members at that time.

Table A.2: Reasons why farmers grow new types of bananas

Reasons for growing "new" bananas (other Tanzanian[a]) (%)	Most important	2nd most important
Declining productivity of indigenous plants	27	15
Experimentation	64	38
Advantages of appearance	5	23
Because others do	5	
For biodiversity		23
Number (%) of households growing/giving reasons for growing	22 (19%)	13

[a] The bananas from other parts of the country were introduced in the Kagera region during the fifties. Bananas from Latin America are introduced by the Kagera Community Development Programme since 1998. Only 4% of the households grows the latter type.

Table A.3: Household characteristics of adopters of new types of bananas versus non-adopters

HH characteristics	Adopters	Non-adopters	Significance of difference
Male-headed households	88 %	69 %	**
Age of household head	44	46	
HH head received some primary	100 %	79 %	***
HH head completed primary	85 %	52 %	***
Value of HH durables (in Tsh)	233710	34014	**

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

Table A.4: Explanation of techniques to maintain banana plant

Technique	Explanation
1. Special way of digging the hole	Hole preparation: when digging the 60 cm deep hole, soil from the top 30 cm should be heaped on one side of the hole and soil from the other 30 cm on the other side. The top soil should be mixed with organic manure (see next technique) and returned to the hole first, in preparation for planting. If this is not enough to fill the hole, top soil from the surrounding areas should be added instead of using the bottom soil.
2. Applying fertiliser/manure	Soil preparation: the best manure to use is farmyard manure from cattle, pigs, goats and chicken, also compost or coffee husk humus can be used. The manure (5 debe or 70 kg) should be thoroughly mixed with the top soil and the hole filled with this mixture should be left undisturbed for minimum 2 weeks.
3. Hot water treatment of the stem before planting	Cleaning of planting material: weevils are mainly located in the roots and corms of the banana plants. Therefore paring is needed (see "paring") and in addition pared suckers and corms can be immersed in hot water, then sterilised and dipped in an appropriate insecticide solution.
4. Dipping stem in insecticide solution	Cleaning of planting material: before planting, dipping the stem in an insecticide solution, used in combination with or without hot water treatment.
5. Mulching 1 meter from stem	Mulching conserves moisture, controls weeds, contributes to soil fertility and reduces soil erosion. But the mulch should be kept away from the base of the plants to prevent superficial root growth.
6. Trench-manuring	Water conservation: the banana plant requires a lot of water and is susceptible to drought. In areas with less then 1000mm of rainfall annually, water conservation methods should be applied. One of the recommended methods of rainwater conservation is trench-manuring. Trenches are dug midway between the banana stools. The bottom of the holes are filled with farm manure and topped up with top soil. Manure absorbs and stores water which the plants can use during the dry season. An alternative to manure is freshly cut banana pseudostem.
7. Paring	Cleaning of planting material: weevils are mainly located in the roots and corms of the banana plants. To reduce the incidence of transferring pests from one infected site to a non-infected one when transplanting suckers one can do the following: remove the roots and pare the corm and then cut off all weevil tunnels.
8. Desuckering (3 plants per stool)	Ideally there should be 3 plants growing on one stool at varying stages of development. Any more suckers deplete the mat of its vital nutrients and provide unnecessary shade.
9. Harvest hygiene	The pseudostem of a harvested banana plant should be cut down at the corm level and soil should be put on the surface to reduce weevil attraction.
10. Weevil trapping	It is not the adult weevils that damage the banana plant but their larvae. Adult weevils are strongly attracted to freshly cut pseudostems and corms so they are ideal for weevil trapping. Split pseudostems are placed facing downwards on the ground on opposite sides of the stem. Continuous cleaning of the trap is necessary.

Source: Mbwana, A.S.S e.a. (1998), "A Guide to Growing Bananas in the Eastern African Highlands", ICIPE

Table A.5: Reasons for not adopting productivity enhancing techniques

Reasons for not using techniques	%
Capital constraint	40
Do not belief in technique	26
Only recently learned	10
Too difficult	7
Labour/time constraint	6
Lack of guidance	4
Technique bad for plants	4
Not (yet) necessary	3
Number of techniques known by farmers but not practised	227

Figure A.1: Technique diffusion

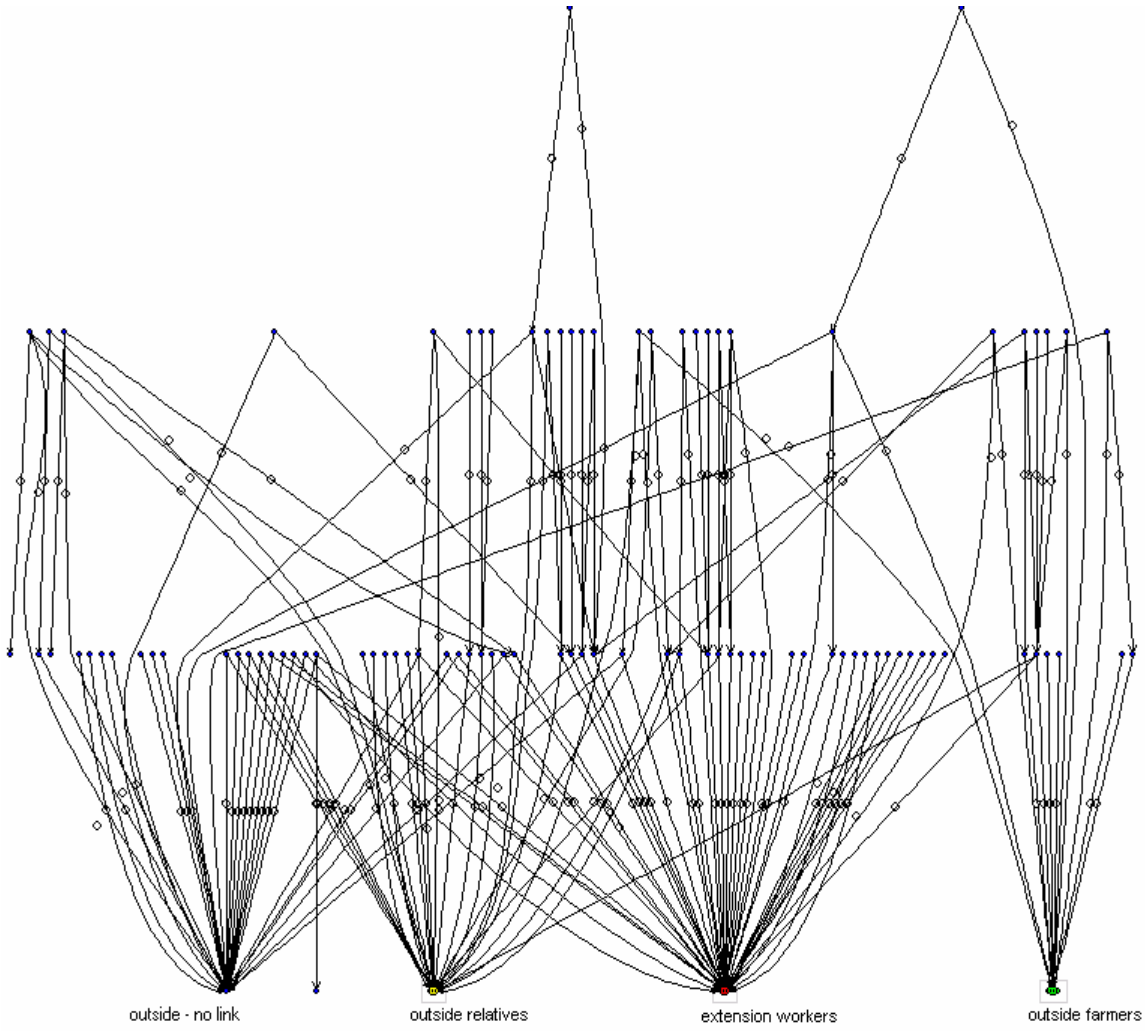


Table A.6: Information links (probit regressions)

Dependent variable: 1 if link exists	Technique learning links (HH)[a]		Hypothetical advice links (IND)[b]	
	Coefficient (1)	Std. err. (2)	Coefficient (3)	Std. err. (4)
(Heads) same age (difference<5)	0.025	0.155	0.079	0.101
(Heads) same gender	0.224	0.151x	0.507	0.096***
Both (heads) lower primary	0.025	0.157	-0.257	0.104***
Both (heads) higher primary	0.046	0.146	-0.006	0.119
Kinship related households	0.582	0.147***	0.331	0.100***
Distance	-0.001	0.000***	-0.000	0.000**
Social insurance members	0.531	0.143***	1.053	0.096***
Belong to same household	-	-	0.283	0.173*
Constant	-2.731	0.179***	-3.427	0.131***
Observations	12210		46017	
Pseudo R^2	0.186		0.251	

[a] Technique learning links are analysed at the household level;
[b] Advice seeking links are analysed at the individual level;
*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

Table A.7: Determinants of information pullers (probit regressions, marginal effects)

Dependent variable:	Learned from outside sources[a]		Learned from formal extension[a]		Advice from all outside sources[b]		Advice from formal extension[b]	
	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.	Coeff.	Std.err.
Gender (HH head)	0.175	0.152	0.104	0.147	0.193	0.079***	0.189	0.080**
Age (head)	0.002	0.004	0.004	0.004	0.001	0.003	0.000	0.003
Standard 1-4 (head)	0.255	0.165x	0.165	0.153	0.191	0.121x	0.206	0.122*
Standard 5-7 (head)	0.181	0.125x	0.193	0.120x	-0.010	0.104	-0.014	0.106
Form 1-4	-		-		0.354	0.094***	0.425	0.087***
Female adults pres.	0.001	0.093	0.033	0.084	-0.054	0.060	-0.079	0.062
Male adults present	-0.075	0.080	-0.020	0.071	-0.015	0.056	0.036	0.058
Land (hectare)	0.219	0.072***	0.149	0.056***	-0.030	0.036	-0.038	0.037
Constant	-2.027	0.700***	-2.275	0.699***	-0.167	0.404	-0.354	0.407
Observations	107		107		174		174	
Pseudo R^2	0.217		0.181		0.087		0.096	

[a] At household level; [b] At individual level; *** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

Figure A.2: Hypothetical advice links

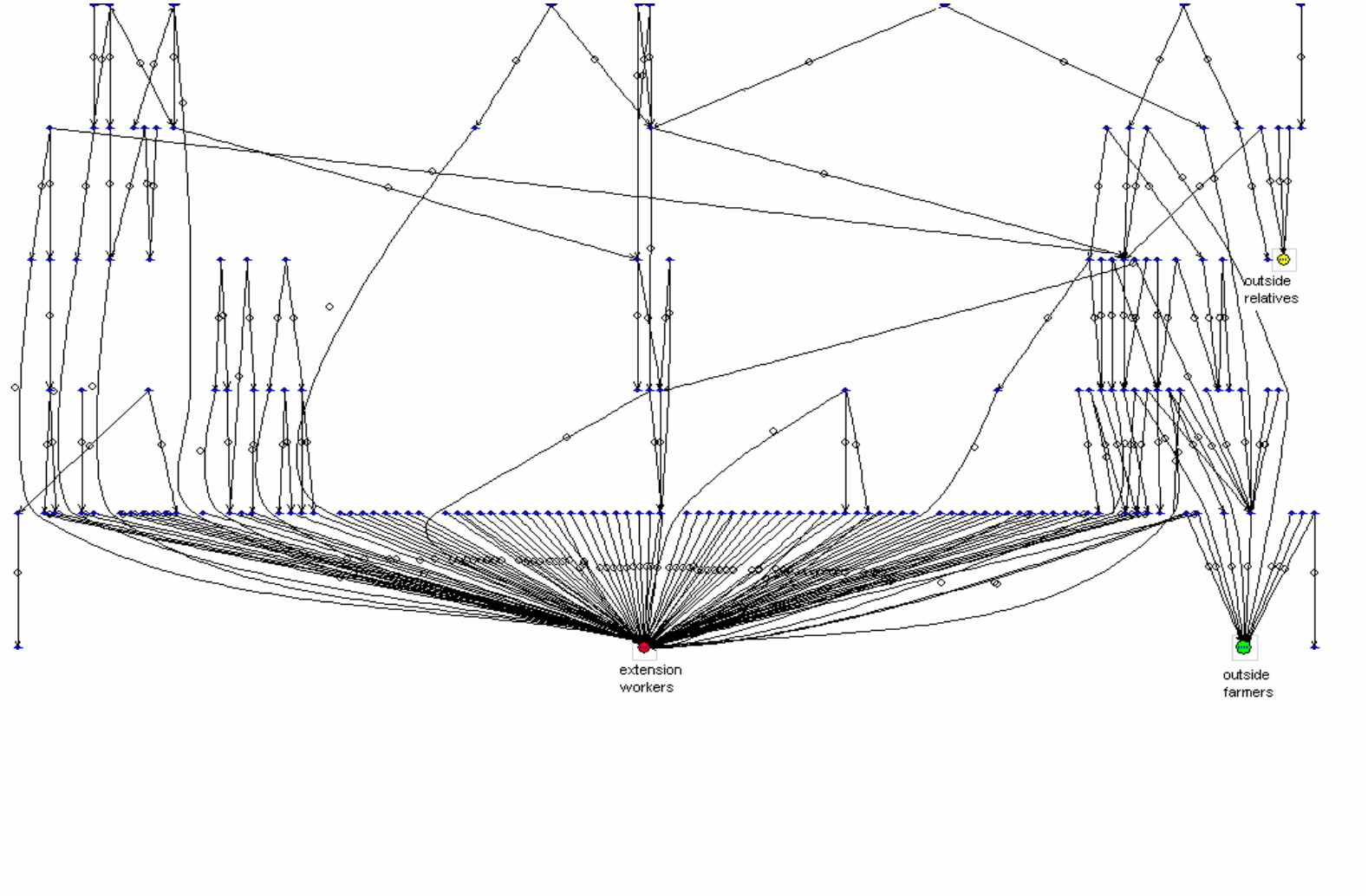


Table A.8: Household adoption, pooled techniques (logit marginal effects)

Dependent variable: use of technique (0/1)	Coeff.	Standard error[a]
Household characteristics		
Total land (hectares)	0.047	0.018***
Adults present	-0.001	0.020
Sex of household head	0.077	0.036**
Head some lower primary	-0.095	0.081
Head completed primary	0.099	0.040***
Age of household head	0.000	0.001
Constant	-2.861	0.612***
Observations	1063	
Pseudo R^2	0.085	

*** significant at 1%; ** significant at 5%

[a] Technique adoption was asked for a series of ten techniques; standard errors are adjusted for household clustering

Table A.9: Differences between lowest versus highest yield quartiles (relative within group)

	Kinship groups		Neighbours		Social support groups	
	<25%	>75%	<25%	>75%	<25%	>75%
Adviser	16	44	16	39	16	38
Teacher	8	15	5	11	11	6
Outside source	54	44	45	58	47	47

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

Table A.10a a: Heckman two-step estimates[a] for endogenous and exogenous social interaction models, all farmers included

Dependent variable: Log(banana output value)	Exogenous group effects assumed zero						Exogenous group effects nonzero					
	(1) Kinship		(2) Neighbours <300m		(3) Social support network		(1) Kinship		(2) Neighbours <300m		(3) Social support network	
	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Grower characteristics												
Log(Kibanja)	0.730	0.792***	0.727	0.217***	0.611	0.196***	0.752	0.183***	0.811	0.196***	0.568	0.217***
Log(Adults)	0.673	0.479	0.662	0.361*	0.890	0.464*	0.518	0.434	0.589	0.318*	0.642	0.381*
Male grower	-0.245	0.290	-0.365	0.277	-0.368	0.277	-0.161	0.271	-0.302	0.253	-0.300	0.268
Some primary school	0.869	0.442**	0.958	0.435**	1.066	0.425***	1.065	0.416***	1.313	0.423***	0.787	0.428*
Completed primary	-0.070	0.381	-0.215	0.366	-0.185	0.378	-0.344	0.392	-0.319	0.340	-0.144	0.355
Log(Age)	0.288	0.546	0.227	0.469	0.363	0.513	0.155	0.504	0.482	0.480	0.341	0.451
Group characteristics												
Log(Yield t-1)	0.178	0.185	-0.486	0.423	0.344	0.180*	0.188	0.173	-0.962	0.429**	0.186	0.176
Nr. lower primary							0.603	0.235***	0.244	0.129*	0.184	0.136
Nr. higher primary							-0.325	0.164**	-0.026	0.066	-0.020	0.126
Nr. male headed							-0.403	0.203**	-0.165	0.109x	-0.081	0.123
Average age							-0.036	0.024x	-0.028	0.054	0.014	0.022
Constant	4.845	3.419	11.069	4.001***	2.617	3.398	6.992	3.481**	13.824	4.538***	3.009	2.782
Mills lambda	-0.013	1.004	0.013	0.848	0.416	1.031	0.015	0.929	0.332	0.847	0.624	0.678
Observations	108		114		116		108		114		116	
Uncensored obs.	87		95		95		87		95		95	
Wald chi²	38.26		37.17		38.32		72.64		60.82		62.42	

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

[a] selection effects not shown in the table

Table A.10b: Heckman two-step estimates[a] for endogenous and exogenous social interaction models, highest yield farmers excluded

Dependent variable: Log(banana output value)	Exogenous group effects assumed zero						Exogenous group effects nonzero					
	(1) Kinship		(2) Neighbours <300m		(3) Social support network		(1) Kinship		(2) Neighbours <300m		(3) Social support network	
	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error	Coeff.	Std. error
Grower characteristics												
Log(Kibanja)	0.779	0.226***	0.677	0.243***	0.611	0.225***	0.758	0.209***	0.758	0.217***	0.605	0.251**
Log(Adults)	0.761	0.539	0.521	0.392	0.870	0.464*	0.593	0.480	0.385	0.339	0.423	0.391
Male grower	-0.082	0.326	-0.272	0.295	-0.323	0.304	-0.053	0.300	-0.187	0.268	-0.274	0.285
Some primary school	0.815	0.486*	0.845	0.456*	1.119	0.435***	1.101	0.449***	1.127	0.433***	0.784	0.428*
Completed primary	-0.001	0.395	-0.278	0.382	-0.248	0.382	-0.375	0.420	-0.406	0.346	-0.227	0.359
Log(Age)	0.402	0.592	0.134	0.484	0.277	0.536	0.163	0.544	0.273	0.486	0.136	0.462
Group characteristics												
Log(Yield t-1)	0.330	0.235	-0.394	0.540	0.412	0.218*	0.317	0.212x	-0.628	0.511	0.241	0.205
Nr. lower primary							0.615	0.266**	0.196	0.135x	0.196	0.137x
Nr. higher primary							-0.335	0.179*	-0.029	0.070	-0.094	0.136
Nr. male headed							-0.403	0.234*	-0.116	0.111	-0.047	0.132
Average age							-0.027	0.027	-0.053	0.059	0.012	0.022
Constant	2.770	4.250	10.796	4.848**	2.214	3.856	5.152	4.117	13.110	5.739**	3.662	3.004
Mills lambda	0.277	1.098	-0.339	0.882	0.548	1.060	0.105	1.001	-0.177	0.861	0.372	0.724
Observations	97		109		106		97		109		106	
Uncensored obs.	76		90		85		76		90		85	
Wald chi²	32.35		29.32		34.77		62.64		51.91		53.91	

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%

[a] selection effects not shown in the table

Table A.11: Effect of group members' frequency of extension contact on individual output (OLS), two highest yield farmers excluded

Dependent variable: log(banana output value in 2000)	Kinship		Neighbours <300m		Social support network	
	Coeff.	Standard Error	Coeff.	Standard Error	Coeff.	Standard Error
Individual characteristics						
Log (Kibanja hectares)	0.697	0.266***	0.765	0.211***	0.530	0.251**
Log(adults)	0.320	0.413	0.375	0.347	0.213	0.404
Male grower (1 if male)	-0.080	0.347	-0.215	0.285	-0.136	0.316
Some primary schooling	1.384	0.547***	1.096	0.433***	0.537	0.474
Completed primary	-0.602	0.479	-0.369	0.363	-0.220	0.383
Log(Age)	0.132	0.500	0.315	0.443	0.048	0.472
Group characteristics						
Total nr. of contact with extension officer in 1999	0.024	0.010**	-0.002	0.007	0.003	0.005
Nr. lower primary	0.656	0.237***	0.183	0.138	0.336	0.152**
Nr. higher primary	-0.306	0.198x	-0.049	0.074	-0.118	0.145
Nr. male headed	-0.461	0.202**	-0.085	0.116	-0.160	0.133
Average age	-0.034	0.031	-0.047	0.064	0.012	0.025
Constant	8.032	2.651***	6.976	3.924*	6.304	2.140***
Observations	66		91		80	
Adjusted R^2	0.262		0.283		0.193	

*** significant at 1%; ** significant at 5%; * significant at 10%; x significant at 15%