KATHOLIEKE UNIVERSITEIT LEUVEN

Faculty of Economics and Applied Economics

Department of Economics

Social Interactions in Growing Bananas

by

Katleen VAN DEN BROECK

Development Economics

Center for Economic Studies
Discussions Paper Series (DPS) 03.12
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

December 2003

DISCUSSION PAPER

Social Interactions in Growing Bananas

Katleen Van den Broeck

December 2003

Abstract

In an environment of strongly decreasing banana productivity, we analyse whether an increase in the average productivity of a reference group a farmer belongs to, has a positive effect on that individual farmer's harvest. The increase in average productivity is supposedly caused by the adoption of productivity enhancing techniques. So we measure the externalities of a productivity increase in one farmer's banana field. For our analysis we have data on three social groups, namely kinship members, neighbours and social insurance group members. We find the strongest social effects within kinship related groups. We do find exogenous social effects between neighbours: there is a positive effect of neighbours' education level. But only within kinship related groups we find the true endogenous effects that produce the social multiplier in banana productivity.

1. Introduction

Technology adoption in farmer's livelihood systems can become an important element of daily survival. In Kagera, in Tanzania, farmers observe a severe productivity decline of the banana, which is their main staple food. Population pressure in the area is high and soil fertility decreases strongly. Moreover, the indigenous banana suffers from many 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. A broad range of literature on the adoption of a new technology exists[1]. It is often found that the probability of adoption is influenced by some individual characteristics of the farmer such as his human capital, the degree of risk aversion, farm size. Additional to a farmer's individual characteristics, a farmer's behaviour can be influenced by other farmers' behaviour. 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).

In our work we will start from the fact that the productivity of the indigenous banana is decreasing and that techniques to mitigate the negative effects are available to the farmers. Farmers can learn about a new technology via different sources. A farmer can directly learn it from extension services or other forms of communication, e.g. pamphlets, books, radio. These farmers are the innovators or the experimenting farmers in the village. But farmers can also learn about a new technology via other farmers. Moreover, the positive or negative attitude of other farmers in a reference group towards the new technology, may influence individual behaviour. We will not analyse behaviour in adopting the techniques directly, but rather look at banana output, which is an indirect measure of farmers' behaviour. Additional to a farmer's characteristics, we will analyse how the output of a reference group can affect output of an individual farmer belonging to that group. We

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

assume that group output will affect individual output in the next agricultural season because of 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 they offer is to use instrumental variables that directly affect outcome of some but not all group members.

Besley and Case (1993) have also addressed some of the problems of using cross-sectional data. They argue that coefficients will be biased due to changing farmer characteristics, e.g. credit availability and a farmer's knowledge about the new technology. Both a farmer's credit situation and his knowledge can change over time. We assume that credit availability does not influence adoption because the new banana varieties are often freely distributed at markets. As for the techniques, some of them do require inputs, but most of them do not. In our paper, a farmer's knowledge about techniques will be represented by the lagged agricultural behaviour of his reference group since we assume he learns about techniques via these other farmers.

A serious 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, 1998) as the reference group. These are large reference groups but 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 insurance 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. On 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 additional important 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.

In the next section we will present an analytical framework for our analysis, mainly drawn from Berger (1985), Udry and Conley (2001) and Manski (1993) and we propose a social

---

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.

interactions test. 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. Section four is the empirical part, where we test for social interactions between farmers. We will look at whether individual banana output is affected by the lagged average banana productivity in a reference group (which was called by Manski the endogenous social effect). We will also correct for individual characteristics and exogenous group characteristics such as education level and age composition of the group (exogenous social effects).

We find the strongest endogenous social effects within kinship related groups. In our sample we do not find social effects within social insurance groups once corrected for exogenous characteristics. This points out the fact that these groups are probably formed on the basis of characteristics that drive not only group formation, but productivity of the group members in the same way. We also find exogenous social effects within neighbour groups. In the last type of groups we find public effects of neighbours' primary education (completed primary or seven years of primary), complementary to the private effects of some years of education (four years). But only within kinship related groups we find the endogenous effects that produce the social multiplier effect. When we look deeper into the direct effects of information obtained by group members, again we find the most positive effects within kinship related groups indicating that at least part of the endogenous social effect is due to information spillovers.

2. Theoretical framework

We use the theoretical framework on Bayesian updating of beliefs provided in chapter four of Berger (1985), the target input model (Udry and Conley, 2001; Foster and Rosenzweig, 1995) and Manski's reflection problem (1993) as the basis for our empirical test of social interactions. 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 in 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 will use the same Bayesian framework, not to learn about the optimal amount of a new input but to learn the actual benefits of adoption. Besley and Case (1994) also analysed the process of updating expectations on profitability for HYV cotton.

In our case, farmer *i* has some prior beliefs about the benefit (production gain minus effort) of applying a technique. In the first period he believes the production gain of applying the technique is not worth the effort or there will be no production gain at all. 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. But each period more information is revealed about the true benefits by other farmers, who do apply the technique, and farmer *i* can update his beliefs using this newly revealed information. When farmer *i* starts believing the benefits are positive, he also applies the technique. We assume this process to be the underlying reason for productivity of group members to be related.

We start from farmer *i*'s observation that his productivity is declining. We assume productivity enhancing techniques are available and he considers applying them to limit the loss he faces, though he is not convinced about their possible benefit. Some of the techniques available were only recently introduced to the farmers by farmer extension officers. To some of the techniques there may be an effort attached. For example, a technique exists to prevent weevils from attacking the banana plant (Annex A.1). 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. Some periods of the day may be better suited to clean these weevil traps, e.g. in the morning.

Assume a farmer in the village knows about the existence of a technique and at *t-1* he has prior beliefs on the benefits (production gain minus effort) of applying the technique, B_{t-1}. Suppose the prior B_{t-1} is a normally distributed function with unknown mean β_{t-1} and known variance τ_{t-1}^2. The true benefit of the technique X is also 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. We assume, rather unrealistically, X to be identically and independently distributed across time and farmers. 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 apply 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 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)

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

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

$\beta_t(x)$ can be positive and farmer *i* can start applying the technique in period *t*. In the presented case farmer *i* learns from all n experimenting farmers in the same way and without any noise. However, both the assumptions of full and of perfect information are rather unrealistic. Udry and Conley (2001) find for pineapple growing farmers in Ghana that information flows via sparse networks rather than through the whole village. Following the authors we will also assume that information is restricted to networks or farmers do not learn from all other farmers in the same way. Also the assumption of full information between farmers will be dropped. Therefore we will adapt (1) and (2) to hold first for imperfect information and then we will further adjust it to correct for different relationships between farmers.

First, we assume that 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):

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

where the farmer attaches more weight to his 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. This may be so for farmers belonging to the same family or the same social insurance network. The variance of the noise is large if the information is revealed by a farmer 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.

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

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

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.

This theory suggests that farmers can 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. This is the type of information the farmer will take into account to update his beliefs on the benefit of a technique. In our empirical work we aim to test whether social learning between farmers in a network exists. In most of the literature the farmers taken to be the farmer's network members are all the farmers of the neighbourhood, the village, the district or the same ethnicity. However, these are all expected interaction groups. Because of data limitations many authors are forced to use these expected interaction groups. Recently, some authors collected information on the exact interaction groups farmers belong too. 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 also have information on actual social groups a farmer is part of, e.g. social insurance 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. The dataset contains detailed network information on who is part of whose social insurance network, on blood bonds between households, on exact geographical distances between households, or on clan members.

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

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

with $x_{i,t}$ 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. As we have showed (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 farmer's 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 network members' fields ($\bar{x}_j$).

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

We assume the process takes the following chronology: in period *t-1* farmer *i* receives information about $x_{j,t-1}$ 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 will try out these techniques on his own farm in period *t-1* and he will obtain increased harvest results in period *t*. So we assume $\bar{x}_{j,t-1}$ to have a positive effect on $x_{i,t}$ if farmers really learn from each other and to have no significant effect otherwise. Foster and Rosenzweig (1995) also investigated the relationship between neighbours' cumulative experiences (or neighbours' profitability of adoption) and farmer's profitability. Udry and Conley's findings (2000) suggest that a farmer increases his fertiliser use after someone in his information network achieves higher than expected profits from increased use of fertiliser, which in turn will increase his own profits.

What we 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-demographic 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 institutional 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 average.

So equation (7) represents the pure endogenous effects model where it is assumed that the coefficients of the exogenous and correlated 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 to 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 is not readily available. It is a typical omitted variables problem which results in biased estimates of the coefficient on $\overline{x}_{j,t-1}$. Therefore we drop the assumption of zero coefficients for the exogenous effects by including them explicitly in the test. The coefficient of the correlated effects is assumed zero since all surveyed farmers live in the same village and face the same environmental factors. 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 (2001) 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[4] 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. Therefore we will have to separate the exogenous group characteristics from the group behaviour itself and our test becomes:

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

If the coefficient $\overline{x}_{j,t-1}$ is non-zero, then there are endogenous social effects. If the coefficient vector of $\overline{Z}_j$ is significantly different from zero, there are exogenous social effects. An alternative method to estimate the endogenous social effects proposed by Manski, could be to apply a two-stage method in the guise of a spatial correlation model:

$$y_i = \beta W_{iN} Y + z_i^{'}\eta + u_i \qquad i = 1,\ldots,N$$

Where $Y = (y_i, i = 1,\ldots, N)$ is an *Nx1* vector of sample realizations y and W_{iN} is a weighting vector, with non-negative components summing to one. $z_i^{'}$ is a vector with individual characteristics. This model can be estimated by maximum likelihood. It was applied, for example, by Case (1992).

4 Brock and Durlauf (2001) show that self-selection may actually facilitate identification. Self-selection may induce the sort of non-linearities that generate identification of the endogenous 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 village 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 average exogenous characteristics.

3. The data: Nyakatoke, a small village in Tanzania

The data we will use for our analysis were collected in a 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 of bananas but the farmers are reluctant to adopt these new kinds. They cling to the old kinds because they claim their taste is better. We do not have exact data on the variance of new banana harvest, where higher variance could possibly explain farmers' reluctance to grow new kinds due to risk aversion. But we assume their variance to be equal to the indigenous banana variance. Unlike with high yielding rice varieties that give much higher output 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. Additional to the introduction of new kinds of bananas, NGOs together with farmer extension centres try to diffuse 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 survey was conducted in Nyakatoke, a small village located at 60 kilometres from Bukoba, the largest town in the region. 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 have information on household characteristics such as land ownership, assets, quality of housing, food consumption, expenditure and household demographics. At the individual level the information collected contains data on income, harvest, time allocation, transfers in kind and in cash, 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 linkages in the village.

Firstly, we will present how farmers perceive the evolution of productivity on their banana farm. In table 3.1 we summarized the answers to the question how the banana harvest of 2000 was to be compared with the banana harvest of ten years ago. We find that a strikingly high number of the respondents that already grew banana in 1990, claim that their banana harvest decreased strongly over the past ten years. This indicates the depth of the problem.

Table 3.1: Respondents observation of banana harvest, 2000 versus 1990

Harvest 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.

Productivity decrease of the banana plant is viewed as one of the major economic problems the village faces[5]. Still, the reaction of the farmers with respect to the adoption of more resistant types of bananas or productivity enhancing techniques is limited (cfr.infra). There seems to be some kind of trade-off between taste and fertility decrease. Farmers prefer the indigenous kind of banana but once the fertility of their field is too low they will start experimenting with new kinds. In the villages located closer to the lake border in the high rainfall zone of Kagera, the soil is already more depleted and most of the farmers have already adopted new kinds of banana. The strong decrease in productivity however has not yet led the farmers to massively experiment with new, more productive kinds of bananas, nor to apply productivity enhancing techniques introduced by farmer extension centres (the techniques that farmers were asked about in the survey are listed in annex table A.1). However limited the reaction of the farmers is, we will devote a separate paragraph both to the adoption of new bananas (3.1) and of productivity enhancing techniques (3.2). In paragraph 3.3 we will describe the process of advice seeking behaviour and in the last paragraph we will analyse the characteristics of farmers who learn from sources out of the village versus those that learn from farmers inside the village. Moreover, we will analyse in depth the intra-village learning or advice links with respect to the relationship between farmers and their teachers or advisers.

---

[5] It is often mentioned as an answer to the following survey question: "Has the situation of the village changed compared to ten years ago? If so, explain."

We are interested in social interactions, i.e. how the behaviour or the exogenous characteristics of the social group a farmer is part of influence his individual behaviour. Our aim is to define a farmer's information network. However, we do not have sufficient direct data on it. Therefore we will use other networks as proxies for the information network. In the following paragraphs we will show that the social networks used (kinship, neighbours and social insurance groups) are good proxies for a farmer's information network. We will focus on three social groups. In order of exogeneity of the group, we will use a farmer's kinship network, his neighbours and his social insurance 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 degree. To form a network of neighbours we have taken 300 metres as the threshold for the household 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 insurance network. It includes all the households one of the household members mentioned he could rely on in case of need and vice versa. Obviously, the social insurance network is endogenously formed. Farmers can choose whom to link up with and will try to form groups with farmers having certain characteristics, e.g. degree of schooling, diversification of income earning activities, productivity, etc. The endogeneity of group formation can bias our test results and lead us to conclude that there is an endogenous social effect, whereas in reality it is an exogenous social effect. We do not face this problem while using the two exogenous networks, kinship networks and neighbours.

3.1. Adoption of new banana varieties

Only 22 percent of the households in our survey village grow other than indigenous kinds of bananas. In Nyakatoke, 19 percent of all households grow some of the bananas originating from other parts of Tanzania and 4 percent grows Latin American bananas[6]. The most important reason for farmers to grow new bananas is not (yet) because of the declining productivity of the indigenous banana[7] but for experimentation in itself (annex

---

6 The bananas from other parts of the country were introduced in the Kagera region during the fifties. The bananas from Latin America are introduced by KCDP (Kagera Community Development Programme) since 1998.

7 Only 27% of the farmers that grow bananas from other parts of Tanzania mention the decline in productivity as the most important reason and 15% as the second most important reason.

table A.2). Thus the adoption of new kinds of bananas is still in its experimental phase and the diffusion process is far from completed. This makes the process ideal to study. Most of the farmers that grow new bananas started growing them since 1999. Only very few grew new bananas before 1999 (14 percent of the households grew new bananas in 1999, five percent in 1998).

Table 3.2: 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 3.2 shows the characteristics of the households that grow new banana varieties. 88 percent of the adopting households are male headed, which is significantly higher than the percentage male-headed households in the non-adopting category. All household heads who experiment with new banana varieties received some primary education and most of them even completed primary education. The adopting households are significantly better educated and wealthier than the non-adopting households. Thus we find that gender of the household head, the level of education and the household wealth strongly influence whether a household will adopt new banana varieties.

Table 3.3 is similar to table 3.2 but instead of describing the individual characteristics of adopters versus non-adopters, we present an overview of their average group characteristics. We find that adopters not only live in households where the heads have reached significantly higher levels of education, but also belong to social groups where household heads on average have higher levels of education. The difference is most significant for having completed primary education in kinship and social insurance networks. For example, in a social insurance network of an adopter of new banana varieties, 66% of the household heads have completed primary education, where this is only 60% of the household heads in the social insurance networks of a non-adopter. Also the gender effect is present in social insurance and neighbours networks. So additional to the private effects of education and gender, also intra-group public effects of these characteristics exist.

Table 3.3: Average household characteristics of social groups of adopters versus non-adopters

Average HH characteristics of:	Kinship network			Neighbours network[a]			Social insurance network		
	Adopt	Non-adopt.	Sig.	Adopt	Non-adopt.	Sig.	Adopt	Non-adopt.	Sig.
Male-headed households	80 %	75 %		76 %	72 %	**	80 %	74 %	*
Age of household head	46	46		45	45		46	47	
Head some primary	90 %	85 %	x	86 %	82 %	**	90 %	85 %	*
Primary completed	70 %	57 %	**	62 %	59 %		66 %	60 %	**
Value of HH durables (Tsh)	78320	93689		103235	75402	*	144100	159894	

[a] the threshold for being "neighbours" is living at less than 300 metres from each other

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

3.2. Use of productivity enhancing techniques

Additional to cultivating new banana varieties farmers can also apply techniques that enhance productivity or try to prevent weevils from attacking the banana plant. In table 3.4 we show the percentages of household heads that know and use a certain productivity enhancing technique. Some techniques are very familiar to the farmers and others are new and not well known, but in both cases there is a large discrepancy between knowing and actually applying the technique. The mostly mentioned reasons for not applying a certain technique were that farmers lack the money to buy the necessary inputs, farmers do not believe in the technique or they have only recently learned it and did not start applying it yet. Some answered the techniques were too difficult (annex tables A.3 and A.4). The capital constraint is most pronounced for those techniques where manure or fertiliser is used. Not many farmers own cows or other animals, which makes manure a very scarce and relatively expensive input. Also chemical fertiliser is both an expensive and not readily available input. Where it concerns techniques that do not require expensive inputs, farmers often do not believe in the effectiveness of the technique.

Table 3.4: Knowledge and use of productivity enhancing techniques*, 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

*for a more detailed explanation of the techniques, see annex A.1

Some of the techniques were already known by the farmers when they started their farm because they saw their parents practising it. Other techniques are very new ones only recently introduced by extension agents because of the urgent need to prevent the banana trees from deteriorating any further. In table 3.5 we present a summary of how the technology is diffused. Broadly, there are two steps. In the first step information enters the village from outside whether it is from farmers living in another village or via contact with farmer extension workers. The first step has a large positive externality: in the second step information on techniques is diffused amongst other farmers in the village.

Table 3.5: Learning and diffusion of techniques

Source	**% of techniques learned from source**	**% of known techniques used**
Information enters village via:		
Farmers outside community	6 (22)*	55
Formal extension (NGO/government/seminars)	48 (189)	47
Other (Self-taught, school)	3 (7)	50
Diffused further via:		
Farmers inside community	38 (150)	43
No specifically identifiable person	7 (29)	32
All techniques learned	100 (397)	44

* number of times answered between brackets

48 percent of all techniques learned by farmers were taught by an extension officer. Another 38 percent of the techniques that farmers know are learned from other farmers living in the same village. 7 percent is learned from farmers inside the village but who can not be clearly specified. They may be dead relatives or a group of farmers. 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 that learn techniques from outside the village have chosen to put some effort into looking for solutions to their agricultural problems 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 well have accidentally learned it while visiting each other or discussing agricultural issues. With respect to these intra-village learning links, we are interested in whether they are determined by the relationship between both farmers. In table 3.6 we present some information on learning clusters.

Table 3.6: Learning links and social groups

Social groups based on	Percentage of households of the same social group if		
	one has learned techniques	no learning link	Significance
Kinship	45	5	***
Clan	38	9	***
Religion	60	35	***
Distance (metres)	249	522	***
Insurance group	55	7	***
Observations (nr)	40	14002	

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

Table 3.6 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. 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 any of these social groups, on each of which we have full information for all respondents, can serve as a proxy 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 the technique). The importance of social groups is also shown by the regressions in annex (table A.5). The left hand sided 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 insurance 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 far more significant explanatory variables, which confirms the descriptive statistics in table 3.6.

3.3. Advice seeking behaviour

Another question asked in the survey was how many times farmers go to others for advice on agricultural issues. We listed some ten different types of persons, such as the community leader, the innovator of the village, the individual's two most important social insurance members, the extension officer, etc. We do not claim we have captured all possible advice sources but they can serve as an indication of the importance of different advice sources. In table 3.7 we show the relative percentages of certain types of persons whom farmers asked for advice in the year 2000. For social insurance network members we only inquired about the two most important members. Asking for all of them would have been very time consuming. Moreover, if any bias exists, it will be an underestimation of the true advice source a whole social insurance network constitutes. So the weight of the social insurance network capturing only the two most important members versus other persons asked will underestimate rather than overestimate the true weight of social network members as advice sources. "Typical advisers" are the farmers mentioned by the respondent as an answer to another question, which was "whom would you go to for advice if there was a problem with your crops". The innovator of the village is the farmer that experiments the most with new crops[8]. The random persons are drawn from all the villagers that earn an own income. They are included to check for significant differences between just any farmer of the village and specific farmers that are part of a person's information network. They serve as a comparison to the other types of persons.

---

[8] The "innovator of the village" was chosen by the enumerators on the basis of how many different kinds of new crops a farmer cultivates, on the magnitude of the area he allocates to them, when he started growing new crops and on how well they grow. It did not require long discussions before we were able to choose the innovator.

Table 3.7: Importance of agricultural advice by source, in 2000

Source	% of all advice asked *	% of all advice asked by gender	
		women	men
Extension officer	10	9	11
Insurance network members	20	16	23
Typical adviser (female)	9	19	1
Typical adviser (male)	21	17	25
Innovator of the village	20	11	29
Random person (female)	0	0	0
Random person (male)	4	7	2

*the percentages do not add up to 100 because there are other sources, not presented in the table

20 percent of the times farmers seek advice about agricultural issues, they go to their social insurance network members. For both men and women their social insurance network members seem to be very important advice sources, at least from those listed. Men have more contact with the village innovator, most probably since the latter person is also a man. Social insurance network members and typical advisers are far more important than the role extension officers play in the advice seeking behaviour of the farmers (10%). Probably we find this result due to the cost attached to advice seeking and the type of advice one is looking for. Visiting a farmer extension worker, who does not reside in the village, at least has a time cost that is larger than the cost of going to another villager. Moreover, if the problem is a rather small one, one will prefer to go to a friend from the village and talk it over while visiting one another. But if one really has a severe problem an extension officer may be perceived to know better what is appropriate to do. We can conclude that the high percentages of advice asked from insurance network members are not just co-incidence since the percentage of advice farmers asked a randomly drawn person from the village is close to zero. It shows that the assumption that a whole village constitutes one information network is untrue, at least in our case. Random villagers are hardly asked for advice[9]. Information will not simply enter the village and be spread to all farmers but it flows via networks (most likely overlapping with other networks). The table supports the idea that the social insurance network is a good proxy for a farmer's intra-village agricultural information network.

So far we only discussed actual information links, learning or advice links as they actually took place. The survey also contained questions on hypothetical advice links. The farmers were asked "whom they would go to for advice if something went wrong with their crops" (cfr. typical adviser). We analysed these answers with respect to social groups the hypothetical advisers belonged to (table 3.8). More or less the same picture emerges as what we found for actual learning links. Of all the farmers mentioned as hypothetical advice sources, 38 percent belong to kin-related households, which is significantly more than the average percentage kin-related households if no hypothetical advice links exist. In annex (table A.6) we find these descriptive statistics strongly confirmed by the regression results. The group dummies (farmers belong to kinship related households, distance between households, farmers belong to households in the same social insurance 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 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 being part of the same household also increases the probability of there being an (ex ante) advice link. Thus farmers will consult other farmers in the household. This may point to the existence of intra-household information flows[10].

Table 3.8: Hypothetical information links and social groups

Social groups based on	Percentage of households of the same social group if		
	one would go for advice to	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%; ** significant at 5%; * significant at 10%; x significant at 15%

9 The sample probability that there is an hypothetical (ex ante) advice link between two villagers is close to zero percent (0.159%).

10 Mostly the direction is from women asking their husbands for advice and not the other way around.

In the remainder of the section we will compare the determinants of intra-village learning and advice links versus outside information sources.

3.4. Outside versus inside sources of learning/advice seeking

In annex figure A.7 we show how all farmers in the village are connected to their typical adviser, the person or persons they would go to for advice if anything went wrong with their crops. Each line represents a link between a farmer and any of his advice sources. The figure in annex A.8 represents households whose head knows about at least one of the banana productivity enhancing techniques linked to whom he learned it from. The difference between both graphs is that the fist graph does not represent what farmers actually do but what they say they would do in case of a serious problem[11] whereas the second graph represents what actually happened. The two graphs differ substantially in the number of layers they consist of and the denseness of the layers. Included in the first layer (lowest in the graph) are those farmers that pull information into the village. The upper layers include farmers who receive their information from these villagers who have sought information outside of the village.

With respect to hypothetical advice seeking (A.7), many farmers in the survey are included in the first layer. They would go to an extension officer or to outside farmers. The layers of intra-village hypothetical advice links are rather thin. There are a number of farmers who are many steps away from an outside advice source (maximum five layers in A.7). In annex (table A.9) we present probit regressions to look deeper into what determines whether a farmer mentions ex ante he would consult an outside source in case of a severe agricultural problem. The regressions mainly reveal that gender and higher education strongly increase the possibility that the farmer mentions an outside advice source. Men have a nearly 50 percent higher probability of contacting an extension officer. Some primary education has a slightly positive effect but some secondary education has a large and significant effect on the ex ante probability of seeking advice from an extension officer.

---

[11] 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)

The second graph is a representation of what actually happened. We asked farmers who learned a productivity enhancing technique from another farmer inside the village, who that farmer was. There are a lot of farmers who have learned some of the techniques they know from outside sources. 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 be it from extension workers or from friends and relatives living in other villages[12]. Then there is a second layer of farmers inside the village tapping information from those that have collected it from outside the village. Unlike in graph A.7 most farmers in the village, who have learned techniques, are linked either directly either indirectly via one other farmer, to an outside learning source. Also for learning new techniques we ran a probit regression (table A.10) to explain which types of farmers belong to the first layer in the information diffusion process. What is very prominent from the regressions on actual behaviour is the insignificance of the gender of the household head contrary to what we found for the ex ante regressions. Some primary education is of weakly significant positive influence to having learned techniques from someone outside the village and completing primary education increases the probability of learning from an extension officer. The amount of land is the most significantly positively correlated factor with the probability of having learned techniques from outside the village sources, meaning it is the largest farmers that usually bring information into the village.

3.5. Conclusion

From the descriptive tables we can conclude that there is a lot of information being transferred between the farmers of a village. We have presented evidence to support our hypothesis that the identified social groups (kinship groups, neighbours and social insurance network members) are good proxies for a farmer's agricultural information network[13]. Therefore they may be used to perform the agricultural social interactions test we proposed.

---

12 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.

13 "*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

Mostly a one-directional relationship exists. There are some key group members who are asked for advice but who get their advice from outside the group or the village. Our results show important externalities of information diffusion within social groups. Possibly, we will have to take that observation into account in our test. The test might not hold for all farmers in the same way. For example, the highest productive farmers in a group may prefer to seek information outside their group. Also those farmers who have outside information sources might not learn from group members. If our social interactions test proves positive, an organisation that aims to inform farmers, would not have to contact or include all the farmers in the programme, but the social multiplier effect will make sure the information spills over to other farmers that are part of the same network. If only those interested are informed, information will be diffused to other members of the social groups that farmer belongs to.

But the remaining question is whether the existence of information flows between farmers implies that their outcomes are related. Therefore, in the next section we will analyse the relation between the average behaviour of the social group a farmer is part of and his own behaviour (in terms of the production of bananas). If endogenous social effects would exist within some of the groups defined, we can conclude that the adoption of productivity enhancing techniques does not only affect the output of the adopting farmers. But whether or not the relation between outcomes can be attributed to technique diffusion and learning is yet another question.

network compared to random clusters drawn from the village. We may safely assume that the social insurance and the communication network significantly overlap.

4. Empirical results

The main part of this section will be the social interactions test. Though first, we will briefly summarize the conclusions from the theoretical and the descriptive parts. Next, we will 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 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 some of the evidence shown 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 possible externalities on individual outcome. Therefore we will correct for the relative productivity position of a farmer within his social group. Subsequently we checked the robustness of our results. To correct for the possibility that outcomes of all farmers in the village simply move in the same direction we will randomly allocate a set of characteristics and productivities to a farmer's social group members. Via bootstrapping we will obtain an estimate of whether the coefficient of the true group members' characteristics is significantly higher. Lastly, we will focus on how average group behaviour directly affects individual behaviour, in terms of knowing or using certain techniques, additional to our analysis of indirect effects via banana outcomes.

In the theoretical section we showed how farmers update their beliefs on the expected benefits of applying a certain technique, based on the experience of other farmers. Farmers update their beliefs based on outcomes of other farmers or on information provided by other farmers. We also showed how the updating mechanism may be dependent on the relationship between farmers. A farmer attaches more weight to information received from "group members" than information received from just any farmer in the village.

In section three we found evidence that information on techniques or advice on agricultural issues flows via certain linkages indeed and does not seem to reach all farmers of the village in the same way. What is apparent from the data is the fact that a farmer's advice seeking and learning behaviour is strongly dependent on the relationship he has with the informative farmer. Farmers who are part of the same social group visit and talk to each frequently[14], thereby providing opportunities to observe closely the proceedings on the other's field and to inform one another on production technologies used. So if a member of a farmer's agricultural information network, which we will proxy by three types of social networks, achieves good results by using a certain technique, he can explain this technique exactly. Afterwards the other farmer can decide to apply it himself. If this is really how the process works, we can assume that improved productivity results on the farm of one of the members of farmer *i*'s social group has positive effects on the outcome of farmer *i*. Consequently, outcome is an indirect measure of farmers' behaviour. Since we have full data on all the members of a farmer's kinship, neighbours and social insurance group members, who are living inside the village, we can test our hypothesis of social interaction.

We will start by testing the pure endogenous effects model, using a log-linearized Cobb-Douglas production function:

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

$Log(Y_{i,t})$ is the logarithm of the value of the banana harvest of farmer *i* at time *t* (measured in Tanzanian Shilling). Z_i is a vector with characteristics of the household farmer *i* belongs to. We will include the number of hectares cultivated with banana trees. It is mostly the whole area of the plot around the house[15] (kibanja), which is the most fertile and most closely located 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

[14] 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.

[15] 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.

next to it[16]. 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 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. 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. 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. 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*, such as gender and age of the banana grower[17]. The education level of farmer *i* will be represented by two dummy variables indicating whether the banana grower has gone through lower primary school[18] or completed primary school (cumulatively). The logarithm is taken of all the continuous variables in the regression and dummies remain in the regression as levels.

$Log(\overline{P}_{-i,t-1})$ is the logarithm of the average banana productivity at time *t-1*, of the households that are part of farmer *i*'s social group, excluding the productivity of farmer *i* himself. We use the average productivity 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

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

17 Often the household head and his spouse are jointly responsible for the banana trees although they do sell separately.

information on banana cultivation from his social network, he will share this information with his fellow banana growing household members. Since we assume the process of diffusion of techniques has not reached an equilibrium situation yet, our data allow us to solve the reflection problem described in Manski (1999) by including lagged productivity 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. If $\delta \neq 0$, we can conclude that social effects exist.

The social effect can be (partly) caused by exogenous characteristics of the whole group that drive productivity. Moreover, in the case of endogenously formed groups such as the social insurance network, the coefficients will be correlated across the network members due to self-selection. Hence, we will allow for exogenous social effects in the next set of regressions:

$$\log(Y_{i,t}) = \alpha + \beta \log(Z_i) + \chi \log(X_i) + \delta \log(\overline{P}_{-i,t-1}) + \varepsilon \log(\overline{X}_{-i}) + u_t \qquad (4.2)$$

where $\overline{X}_{-i}$ is a vector of average group characteristics, such as age composition of the group and the average level of education within the group. If $\delta \neq 0$ there are true endogenous social effects. If $\varepsilon \neq 0$, exogenous social effects exist.

Averages of the variables used in the regressions are presented in table 4.1. Of all the individuals in the village who grow bananas, 81 percent mentioned selling bananas as an income earning activity. For them, we have information on total harvest value, both what was sold and what was consumed by the household. But unfortunately we do not have information on consumed harvest values for those households that did not mention banana cultivation as a source of cash income. The reason is that the survey aimed to capture only the income earning activities individuals where engaged in and no information was gathered on harvest values of those crops grown only for household consumption. Although at least one person in each household grows bananas some observations could not be used due to the formulation of the question and missing harvest values. Thus we

---

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

will use a two-step Heckman selection model to correct for the sample selection problem. We will include a first step where an individual has to decide whether to sell bananas or not. As first-step selection variables we will 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 productivity of the banana grower.

To value the harvest, we used prices of 2000 for both years[19], in order to avoid price effects and to capture pure productivity effects. The average output of banana selling individuals was much higher in 2000 than it was in 1999. The reasons for this increase we can only guess. Possibly, there are more trees on the field or more cooking banana trees relative to beer bananas (the latter are much cheaper).

Regarding some characteristics of the banana growers, we see that the gender ratio is fairly equal. As many women as men grow bananas. 79 percent reached the fourth year of primary school whereas only 57 percent completed primary school. Household land is on average 1,3 hectares, half of which is cultivated with bananas. With respect to social groups, there are some households that do not have any kin related households in the village. The number of kin-related households who also sell bananas is around five. The number in the social insurance group is somewhat higher. Neighbours form the largest group. Obviously 300 metres is an arbitrary measure. We chose it because on the one hand, it is below the average distance in the village (523 metres). On the other hand if we were to opt for a lower threshold many observations would be dropped from our analysis.

19 To value the harvest of 1999, we used the average price of a bunch of bananas in 2000 (average of five rounds).

Table 4.1: Description of regression variables

Variables*	Obs	Average	Min	Max
Banana sellers in 2000 (% of banana growers)	117	81	0	1
Banana sellers in 1999 (% of banana growers)	117	67	0	1
Average price of a bunch of cooking bananas	5	1140	1000	1500
Average price of a bunch of beer bananas	5	330	250	400
Output value in 2000 (in 2000 Tsh)	95	6841	71	47878
Output value in 1999 (in 2000 Tsh-average)	81	3351	55	31350
Of all banana growers (95):				
Male growers (%)	95	52	0	1
Age	95	44	20	89
Some primary education (% reached standard 1 to 4)	95	79	0	1
Completed primary education (% standard 5 to 7)	95	57	0	1
Of all households with at least one banana grower (81):				
Kibanja (hectare)	81	0.6	0.03	2.0
Land (hectare)	78	1.3	0.08	8.3
Fertility of the kibanja higher than average (%)	81	19	0	1
Durables (value in Tsh)**	81	90146	400	4357500
Adults present in the household (> 15 years)	81	3	1	8
Children present in the household (<= 15 years)	81	2	0	7
Social groups (average nr of persons in group):				
Kinship members	88	8	1	20
Banana selling kinship members 1999	88	5	1	13
Neighbours (living at less than 300 metres)	95	32	7	54
Banana selling neighbours 1999	95	21	3	34
Social insurance network members	95	10	1	32
Banana selling social insurance network members 1999	95	7	1	25

*values are expressed in Tanzanian Shilling (1 US$=+/-800Tsh)
**in the regressions we used an index of durables rather than the value of the household durables

Tables 4.2a and 4.2b are organised as follows. Table 4.2a gives the Cobb-Douglas estimation results for the case where exogenous social effects are assumed to be non-existent. The first three columns show the results for all farmers in the village, the last three give the results when we exclude the two relatively most productive farmers in each group (the intermediate case when only the first most productive farmers are excluded is presented in annex). We exclude the two most productive farmers in the group[20] because they will be teachers or advisers rather than learners. This assumption is confirmed by our data. In annex table A.11 we find that in the group of the 25 percent most productive farmers there are significantly more agricultural advisers and technique teachers than in the lowest productive groups. Including them in the regressions would underestimate any social effects. This is only possible due to the nature of our data. Since we know the exact group composition and have information on productivity of all members we are able to exclude those most productive farmers. Consequently we can grasp more details of the process. In table 4.2b we relax the assumption of zero coefficients for the exogenous effects and allow them to exist. We will use three types of social groups. In order of exogeneity of group formation these are kinship related groups (completely exogenous), neighbours[21] (relatively exogenous) and social insurance groups (endogenously formed).

Without including any social effects (Annex table A.12) the most significant determinants of banana output value are the magnitude of the kibanja (significant at one percent), the number of adults (significant at ten percent) and having received some primary education (significant at five percent). In the selection, an individual's age and available household labour influence participation positively, whereas the wealth of the household has a negative effect on the probability of selling bananas. From the regressions in tables 4.2a and 4.2b we find that the positive effects of having a larger kibanja, more adults present in the household and having received some primary education are consistent throughout all the regression specifications.

The private effects of having a larger kibanja, higher availability of labour and having run through the first four years of primary school have more or less the same effects and

---

20 Note that the most productive farmers excluded from the group are therefore not the same as the farmers in the first layer of technique learners. We exclude the two most productive farmers since farmers learn from and go for advice to the most productive farmers in their group. The relative productivity position in the reference group was taken from the banana productivity in 1999 (often this position was the same in 2000).

21 Living at less than 300 metres from each other. Neighbours are fairly exogenous since the plot of land where the house is built is mostly inherited from father to sons.

significance in all model specifications. The social effects however, are quite different depending on both the group and the model we analyse.

In the pure endogenous model (4.2a) for all farmers in the group included we find social effects within social insurance groups but not in kinship related groups or neighbours. We have tried out several thresholds of being neighbours, between 50 and 1000 metres, but in none of these we found social effects. When we exclude the most productive farmers in each group (annex A.13a) the coefficients of kinship groups and social insurance groups become larger. Finally, when we exclude the two most productive farmers, we find significant social effects for kinship related and social insurance groups. So we do find evidence to support our hypothesis that there is a direction to the social effects going from the more productive to the less productive group members.

Table 4.2a: Heckman two-step estimates[a] for social interactions, respective social groups

Dep. Variable: Log(banana value)	All farmers included						Two most productive farmers per group excluded					
	(1) Kinship		(2) Neighbours <300m		(3) Social Insurance		(4) Kinship		(5) Neighbours <300m		(6) Social Insurance	
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Log(kibanja)	0.730	***	0.727	***	0.611	***	0.621	**	0.681	***	0.618	***
Log(adults)	0.673	x	0.662	*	0.890	**	1.279	*	0.532	*	0.977	*
Male grower	-0.245		-0.365		-0.368		0.162		-0.300		-0.092	
Some primary educ.	0.869	**	0.958	**	1.066	***	0.913	x	0.914	**	1.085	**
Completed primary	-0.070		-0.215		-0.185		0.333		-0.264		-0.145	
Log(age)	0.288		0.227		0.363		0.998		0.087		0.426	
Group averages:												
Log(productivity t-1)	0.178		-0.486		0.344	**	0.603	*	-0.380		0.456	*
Constant	4.845	x	11.069		2.617		-3.328		10.788	**	0.828	
Mills lambda	-0.013		0.013		0.416		1.421		-0.307		1.129	
Observations	108		114		116		79		107		97	
Uncensored obs.	87		95		95		58		88		76	
Wald chi²	38.22		37.71		38.32		28.37		29.32		29.78	

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

[a] selection effects not shown in the table

Table 4.2b: Heckman two-step estimates[a] for endogenous versus exogenous social interactions, respective social groups

Dep. Variable: Log(banana value)	All farmers included						Two most productive farmers per group excluded					
	(1) Kinship		(2) Neighbours <300m		(3) Social Insurance		(4) Kinship		(5) Neighbours <300m		(6) Social Insurance	
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Log(kibanja)	0.789	***	0.676	***	0.586	***	0.643	**	0.581	***	0.567	**
Log(adults)	0.630	x	0.863	**	0.936	***	1.141	*	0.760	*	0.681	*
Male grower	-0.265		-0.316		-0.357		0.017		-202		-0.099	
Some primary educ.	0.923	**	1.034	**	1.038	**	1.094	*	0.965	**	0.909	*
Completed primary	-0.158		-0.355		-0.158		0.088		-0.343		-0.156	
Log(age)	0.210		0.203		0.409		0.771		0.061		0.179	
Group averages:												
Log(productivity t-1)	0.151		-0.582		0.297		0.536	*	-0.774		0.272	
Head some primary	0.620		-1.171		0.531		0.093		-1.803		1.591	
Head completed prim	-2.037	*	2.595	*	-0.033		-1.864		2.753	*	-1.569	
Log(age head)	-1.997	*	-1.772		0.172		-0.947		-3.209		-0.071	
Constant	13.734	**	18.074	x	1.706		3.035		26.245	**	3.822	
Mills lambda	-0.022		0.007		0.519		1.225		-0.212		0.418	
Observations	108		114		116		79		107		97	
Uncensored obs.	87		95		95		58		88		76	
Wald chi²	52.78		44.85		48.72		34.12		37.76		38.88	

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

[a] selection effects not shown in the table

Thus we find evidence that social capital truly is capital in the sense that it contributes to productivity (Narayan and Pritchett, 1999). Although we can conclude from table 4.2a that there are social effects within kinship and insurance related groups we can not yet draw any conclusions on the social multiplier effect (or the existence of endogenous social effects). It may well be that outcomes of the farmers in a group are related simply because they have the same exogenous characteristics. Especially in endogenously formed groups such as social insurance groups, the probability exists that the groups are formed on the basis of agricultural productivity of the members or on other factors that also drive productivity such as education. Therefore, instead of an endogenous social effect, it may be that all members have the same exogenous characteristics that drive group formation and productivity at the same time.

To correct for this, we include exogenous average group characteristics (the individual's characteristics again excluded from his group average), such as the percentage of household heads of the group members having received lower primary and higher primary education and the average age of the household heads. From table 4.2b we find that there are no endogenous social effects in any of our groups when we include all farmers in the group. Once we allow for a directional relationship between members of a social group and we drop the two most productive farmers from the group, we obtain the following results. Column four in table 4.2b shows that endogenous social effects exist within kin related groups. From column five we find that there are positive effects of neighbours' education on individual outcome. Apparently, there are public effects of higher primary education between neighbours. This effect is comparable to the positive intra-household externality of a literate household member described in Basu and Foster (1998). It is striking that once corrected for exogenous characteristics (though they do not seem significant) the social effect within insurance groups disappears. So it is indeed the exogenous characteristics that drive group formation appear to affect productivity too. When we do not correct for exogenous characteristics the social effect in this type of groups appears misleadingly large.

As a robustness test[22] we bootstrapped the results to correct for a possible positive trend in the village affecting all farmers in the same way. The question was whether the relation between lagged group member's productivity and own harvest value only holds for kinship and social insurance group members. Possibly the relation also exists if we were to include randomly chosen lagged group productivity averages. We constructed the test as follows. In the file with household characteristics (and lagged productivity) we randomized the household numbers. After linking this file to the file containing the individuals and all their kinship, neighbour and social insurance households, these households have now randomly attributed characteristics and lagged productivity. Thus it will produce random lagged average group productivities for the Heckman regressions explaining individual banana harvest values. In figures 4.3a, 4.3b and 4.3c we find the results of 1000 random matches for kinship, neighbours and social insurance groups respectively. The bootstrap test confirms our results. For kinship and social insurance groups with randomly attributed characteristics, the coefficient of lagged average group productivity was lower than the coefficient with group members' true characteristics in at least 90 percent of the drawings.

To summarize the results of the regressions, we found that there are endogenous social effects between kinship related farmers but they work from the more to the less productive members. Thus kinship members only learn from their most productive family members. With respect to neighbours, we do not find any endogenous social effects but there are exogenous social effects of higher primary education. So additional to the private effects of lower primary education there seem to be public effects of higher primary education. So farmers living in neighbourhoods with a relatively high concentration of households with a

---

22 Another robustness check we performed was a fixed effects estimation to correct for farmer heterogeneity 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 that change over time. Though we have to remark that the number of households where adults present changed was only marginal. We ran an Ordinary Least Squares regression instead of using the Heckman selection model. We assume the OLS coefficients to be consistent estimates since the coefficient on Mills lambda proved insignificant so there is no selection problem. In the fixed effects regressions we find positive but insignificant coefficients for lagged group productivity in all three groups. This may be due to the fact that the lags are rather short in order to measure any improved productivity effects in growing bananas. Banana plants need approximately one year before they are fully grown, so three or six months lags are not the appropriate period for measuring learning effects. What we measure by the change in outcome might simply be an approximation of the measurement error and not the change in output. Brock and Durlauf (2001) also 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.

head who completed primary education, obtain higher outcomes. With respect to social effects within insurance groups we can conclude the following. Since these groups are endogenously formed productive farmers may try to link up with other farmers of more or less the same productivity (driven by exogenous characteristics common to all farmers in the group, which have determined the group formation). Therefore, in first instance a social effect appeared to exist. However, this effect disappeared as soon as we corrected for exogenous characteristics.

Figure 4.3a: Bootstrap results for kinship groups

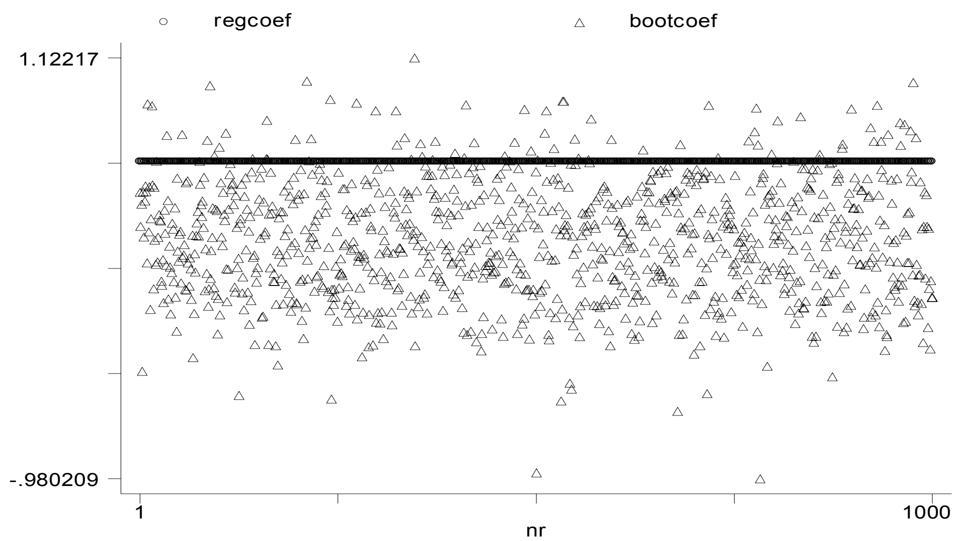


Figure 4.3b: Bootstrap results for neighbour groups

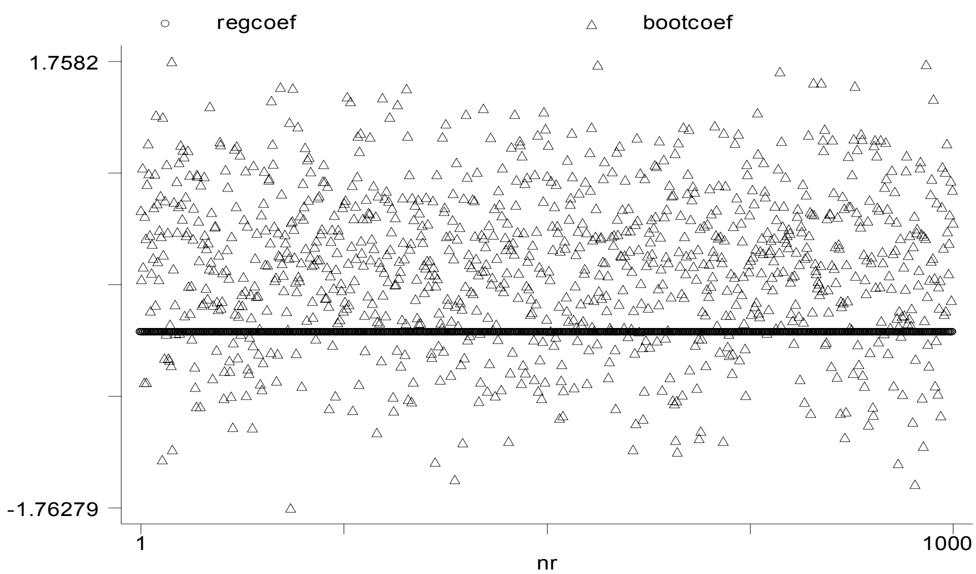
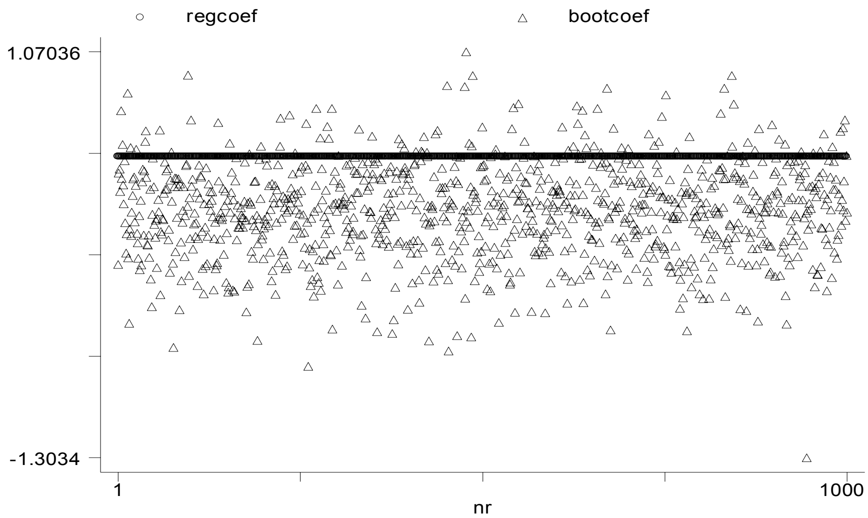


Figure 4.3c: Bootstrap results for social insurance groups



A question that remains is why the endogenous social effects that generate the social multiplier only seem to exist in kinship related groups. Possibly the answer lies in the fact that, before one can gain from another farmer's knowledge, information about the production technology has to be passed on very meticulously (the variance of the noise in transmitting information is small). Simply observing what happens on another farm is not sufficient, but farmers need to know the exact way of how the good result was obtained. Presumably kinship related farmers, e.g. parents and siblings take more time to explain the production technology, than would neighbours or social insurance network members do.

To analyse this hypothesis a bit further, we will look into the direct individual behavioural effects of group members' average behaviour. More specifically, we will analyse whether the average behaviour of the reference group with respect to using techniques or seeking advice positively affects an individual farmers' technique using behaviour or outcome respectively. In table 4.4 we show the results of the probit regressions explaining whether the household head uses (at least one of the) productivity enhancing techniques. In the next table (4.5) we include the average group advice seeking behaviour as an explanatory variable for the banana harvest value. Unlike in table 4.2 we use simple ordinary least squares for the regressions because from 4.2 we found that there was no selection bias.

The regressions in table 4.4 show how the number of techniques used by farmers in a certain social group affects an individual group member's use of at least one of the techniques available. The questions on use of techniques were only asked at household level, so it is mostly the household head who responded. Other than technique use in the group we included the same household and individual variables as in the previous regressions. At the household level we found that the size of the farm is a significantly positive determinant to the using techniques. Education effects only exist for farmers who completed primary school. Apparently there is some education threshold at more than four years of schooling for education to have an impact on a farmer's decision to use productivity enhancing techniques. Supportive to our hypothesis that family members transmit information better, we find the effect of the number of techniques used by group members to be positive and significant only for kinship related groups. What we find is a pure endogenous social effect since we corrected for exogenous group characteristics.

Further, we analysed the effects of information obtained by group members on the production value of an individual farmer (4.5). For information we used first the number of group members having an out-of-village advice source. Alternatively, we used the frequency of group members consulting an extension officer in period *t-1*. The regressions are similar to those in table 4.2, except that we replaced lagged average group productivity by the information obtained within a social group in order to capture the effects of information spreading explicitly. We ran the regression for all farmers (without excluding the most productive ones).

With respect to individual and household level variables the effects are consistent with the results shown in table 4.2. Regarding the advice seeking behaviour of the group, we find that for all three groups endogenous social effects exist: the more group members have an outside advice source the higher is individual banana outcome. Where it concerns the frequency of consulting an extension officer, we only find positive and significant effects within kinship related groups.

Table 4.4: Determinants of household heads' use of at least one productivity enhancing technique

At least one technique used	Kinship		Neighbours		Social insurance network	
	Coefficient	Significance	Coefficient	Significance	Coefficient	Significance
Total land (hectares)	0.831	***	1.046	***	0.748	***
Adults present	0.163		0.189		0.208	
Sex of household head	-0.923	*	-0.442		-0.447	
Age of household head	-0.016		-0.012		-0.011	
Head received lower primary	0.016		-0.394		-0.611	
Head completed primary	0.907	**	1.182	***	1.177	***
Nr of techniques used by group members	0.225	*	-0.154	**	0.043	
% lower primary in group	3.912	**	-3.940	x	2.979	***
% higher primary in group	-2.079	*	0.724		-0.341	
Average age in group	-0.049	*	-0.088		-0.020	
Constant	-0.152		6.386	x	-1.737	
Observations	85		93		93	
Pseudo R^2	0.304		0.296		0.317	

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

Table 4.5: Effect of group members advice sources/frequency of extension contact on individual production values (OLS)

Dependent variable: log(monthly banana harvest value)	Kinship				Neighbours				Social insurance network			
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Log (Kibanja hectares)	0.791	***	0.834	***	0.704	***	0.607	***	0.609	***	0.620	***
Log(adults15-65)	0.764	**	0.465		0.653	**	0.841	***	0.436		0.723	**
Sex	-0.228		-0.272		-0.176		-0.275		-0.309		-0.283	
Received some primary	0.912	**	1.092	**	1.237	***	0.963	**	0.820	*	0.839	*
Completed primary	-0.185		-0.352		-0.472		-0.310		-0.287		-0.268	
Log(age)	-0.001		-0.001		0.002		0.001		0.002		0.001	
Group members having outside advice sources	0.095	*			0.074	***			0.105	***		
Group contact with extension officer in 1999			0.015	*			-0.009	x			0.005	
% lower primary (group)	1.058		0.598		2.273		-2.635		1.006		1.135	
% higher primary	-1.913	*	-1.773	*	0.758		1.691		0.443		0.329	
Average age	-0.023		-0.035		-0.009		-0.080		0.024		0.015	
Constant	8.100	***	9.613	***	3.396		12.143	***	4.641	***	5.425	***
Observations	84		84		91		91		91		91	
Adjusted R^2	0.298		0.298		0.302		0.215		0.288		0.203	

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

Concluding, we found social effects, either endogenous or exogenous ones, within all three groups but the endogenous social effects producing the social multiplier effect, were only found within kinship related groups. We hypothesized that this is due to better transmission of information (lower variance of the noise in transmitting information) which was acquired by family members. Supportive to this hypothesis we do find that information spillovers (use of techniques, individual outcome effects of group advice seeking behaviour) are most prominent in kinship related groups. However, the individual effect of group information appears small relative to the total social effect we found in the first sets of regressions (table 4.2). This leads us to conclude that there are other mechanisms besides information spillovers that generate the endogenous social effect.

5. Conclusion

We analysed whether social groups affect an individual farmer's production behaviour. The productivity of the East African Highland banana has been decreasing since many years due to the declining fertility of the soil, banana diseases and weevil attacks. In Kagera, west of Lake Victoria, the banana is the main staple food and an important cash crop. Therefore it is crucial both for food security and for cash income that this trend be reversed. Farmer extension centres and NGOs alike 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 seem to be interested in learning new techniques and even less so in actually applying them. In our work we aimed to provide evidence that possibly there are externalities of even few farmers learning and using techniques. The true diffusion of techniques is not limited to those farmers who have extension contact. Firstly, farmers will inform other farmers about the techniques so knowledge will be spread. Secondly, if farmers succeed in increasing their productivity, this will have positive effects on the productivity of other farmers in their social group.

We have used a Bayesian framework for updating of a farmer's beliefs about the benefits of a certain technique. 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 whom he knows better, in our case who belong to a social group the farmer is part of. This lead to a social interactions test of the Manski type: individual achievement varies with 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 productivity instead of contemporaneous behaviour to solve the endogeneity 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 group composition and we have data on 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 insurance network members. We found that endogenous social effects exist within kinship groups and exogenous effects within neighbour groups. The latter effect was due to positive effects of higher primary education of neighbours. Following our theory we claim that the endogenous effects within families are at least partly caused by information transmission between farmers. That the effects are largest within families is probably due to more meticulous information transmission. Thus to isolate the effects of information we used specific information variables such as the number of techniques used by group members, the number of group members having an outside (out-of-village) advice source, the frequency of extension contact by group members. Our hypothesis was confirmed by the data: on average the most prominent effects of information spillovers were found amongst kinship related groups.

Thus we can conclude that there is a lot of transmission of information within villages, the effects on productivity being largest for kinship related groups. Those who learn to change their production technology from outside sources will further diffuse their knowledge within their social groups. For positive group productivity effects it seems better to target farmers belonging to different kinship groups. Although the latter variable is a strong determinant of social insurance group formation and even kinship related groups tend to live in the same geographical area, it is not through living close together or through forming groups that endogenous social effects in growing bananas come 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.

Greene, W.H., (2000), "*Econometric Analysis*", Prentice Hall International.

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.

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

Reasons for growing (%)	Growing other Tanzanian	
	1st imp	2nd imp
Declining productivity of indigenous plants	27	15
Experimentation	64	38
Advantages of appearance	5	23
Because others do	5	
For biodiversity		23
Number of farmers growing/giving reasons for growing	22	13

Table A.3: Reasons why productivity enhancing techniques are not used by farmers who do know the technique (part 1)

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

Table A.4: Reasons why productivity enhancing techniques are not used by farmers who do know the technique (part 2), by technique

Reasons for not using techniques (%)	**T1***	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**	**T8**	**T9**	**T10**
Capital constraint	62	85	18	27	9	79	30	12	4	
Do not belief in technique	16	2	18	18	18	11	60	24	70	44
Not (yet) necessary	3	6			9					6
Only recently learned	5	2	47	18	27		10	12	4	6
Labour/time constraint	5		6			5			4	28
Too difficult	8					5		41	7	9
Technique bad for plants		1		18	18			12	11	
Lack of guidance			12	18	18					6

* cfr Table A.1

Table A.5: Probit regression explaining the existence of a learning link between two households

	Coefficient	Significance
Heads same age group (difference max. 5 years)	0.025	
Heads of same gender	0.224	x
Both heads lower primary	0.025	
Both heads higher primary	0.046	
Kinship related households	0.582	***
Distance	-0.001	***
Social insurance members	0.531	***
Constant	-2.731	***
Observations	12210	
Pseudo R^2	0.186	

Table A.6: Probit regression explaining the existence of an ex ante advice link between two respondents

	Coefficient	Significance
Same age group (difference max. 5 years)	0.143	
Same gender	0.527	***
Both received lower primary	-0.271	**
Both received higher primary	-0.003	
Belong to same household	0.345	**
Belong to kinship related households	0.325	***
Distance	-0.000	***
Belong to households in same social insurance network	1.029	***
Constant	-3.475	***
Observations	46016	
Pseudo R^2	0.254	

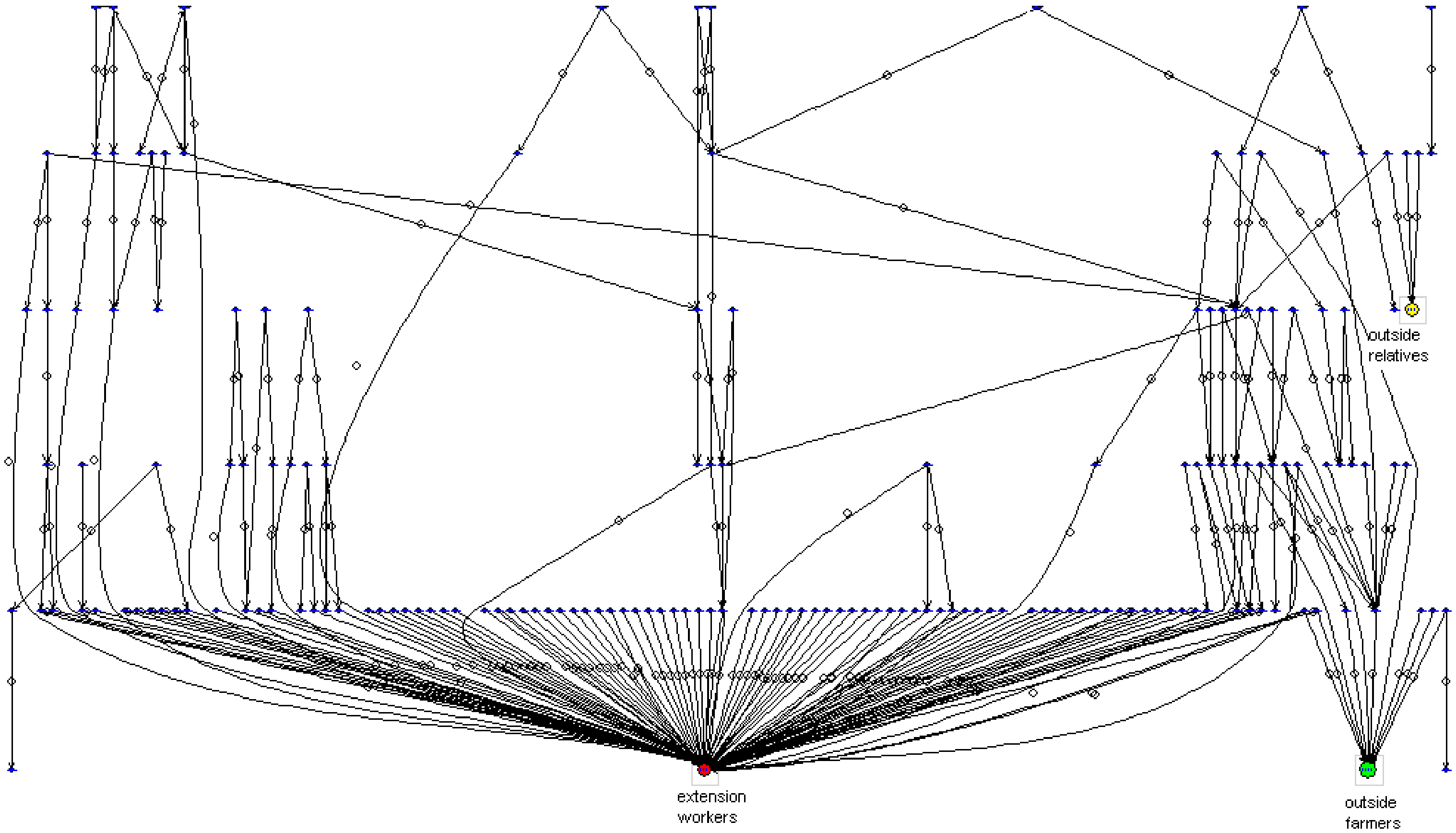


Figure A.7: Hypothetical advice links

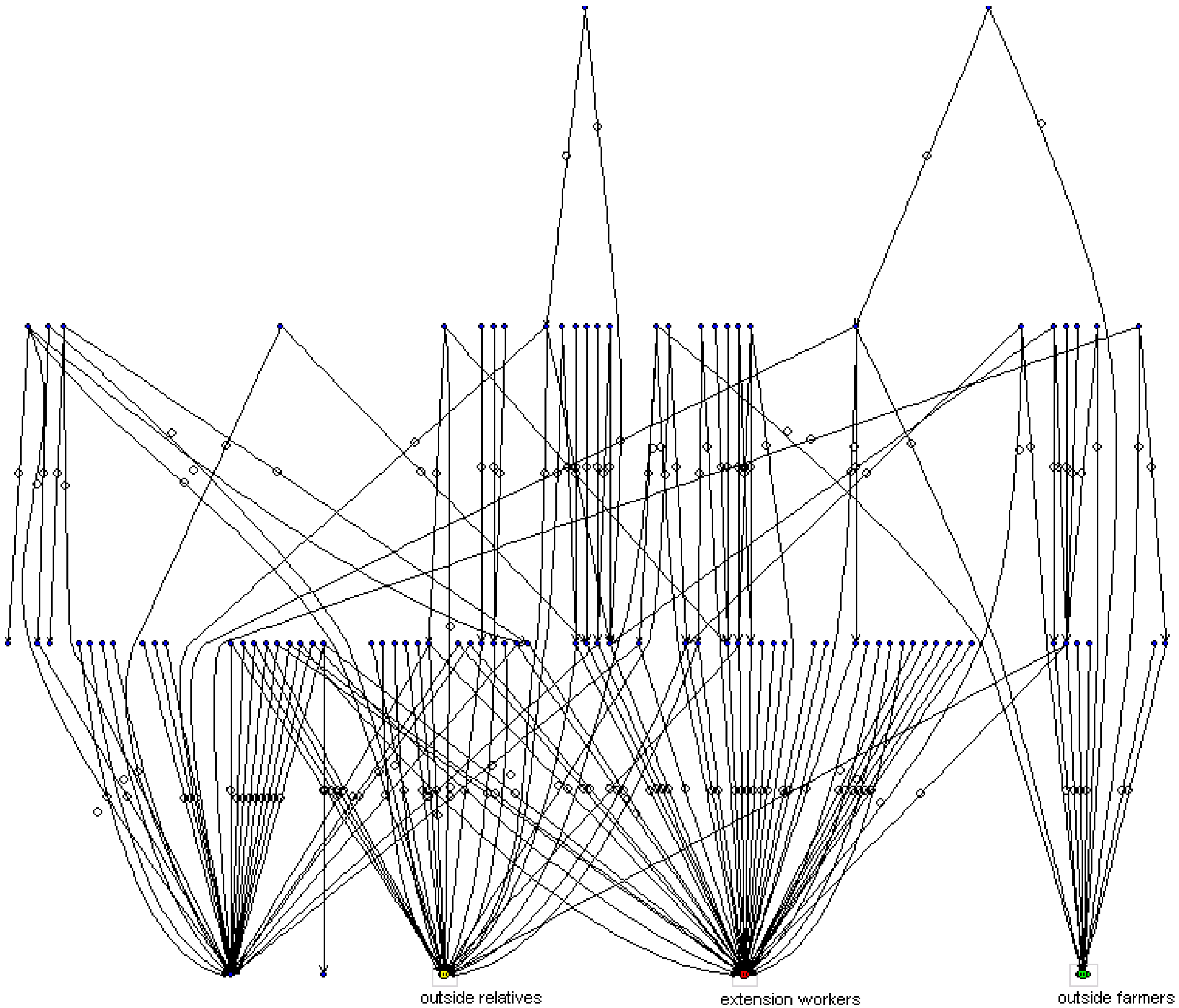


Figure A.8: Actual advice links (techniques learned from)

Table A.9: Probit regressions explaining advice seeking behaviour (ex ante)

Dependent variable (ex ante):	Advice from all outside sources Coeff.	Sig.	Advice from formal extension Coeff.	Sig.
Gender	0.521	**	0.484	**
Age	0.001		-0.000	
Standard 1-4	0.417		0.457	x
Standard 5-7	-0.048		-0.068	
Form 1-4	1.254	**	1.485	***
Female adults present	-0.201		-0.259	*
Male adults present	-0.055		0.114	
Land (hectare)	-0.037		-0.067	
Constant	-0.026		-0.233	
Observations	173		173	
Pseudo R^2	0.087		0.098	

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

Table A.10: Probit regressions explaining outside sources of learning

Dependent variable:	Learned from outside sources Coeff.	Signif.	Learned from formal extension Coeff.	Signif.
Gender of HH head	0.399		0.190	
Age of HH head	0.004		0.009	
Standard 1-4 (head)	0.725	x	0.519	
Standard 5-7 (head)	0.456		0.507	x
Female adults present	-0.030		0.049	
Male adults present	-0.087		0.075	
Land (hectare)	0.529	***	0.373	***
Constant	-2.020	***	-2.244	***
Observations	107		107	
Pseudo R^2	0.212		0.180	

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

Table A.11: Differences between least versus most productive quartiles (relative within group)

	Kin related groups			Neighbour groups			Social ins. groups		
	Lowest 25%	Over 25%	Single-sided t	Lowest 25%	Over 25%	Single-sided t	Lowest 25%	Over 25%	Single-sided t
Gender	51%	53%		45%	51%		45%	52%	
St1-4	73%	81%		63%	83%	***	61%	85%	***
St5-7	59%	56%		50%	58%		42%	62%	**
Age	38	45	**	42	44		43	43	
Adviser	16%	36%	**	16%	38%	***	16%	37%	***
Teacher	8%	15%	x	5%	16%	**	11%	13%	
F adults	1.3	1.6	**	1.3	1.6	**	1.4	1.6	
M adults	1.1	1.6	***	1.1	1.6	***	1.1	1.6	***
Land (ha)	0.9	1.4	**	1.0	1.3	*	0.9	1.4	**
Kibanja	0.5	0.6	x	0.5	0.6		0.5	0.6	*
Cattle	0.4	0.7		0.2	0.8	x	0.2	0.7	
Durables	24189	106594		29193	102544		24629	102439	

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

Table A.12: Heckman two-step estimates Cobb-Douglas

	Coefficient	Signif.
Log(kibanja)	0.672	***
Log(adults)	0.793	*
Male grower	-0.325	
Some primary educ.	0.923	**
Completed primary	-0.183	
Log(age)	0.288	
Constant	6.255	***
Selection:		
Children	-0.012	
Durables (index)	-0.014	**
Log(kibanja)	0.224	
Log(adults)	1.455	***
Male grower	0.081	
Some primary educ.	0.384	
Completed primary	0.284	
Log(age)	1.173	**
Constant	-4.440	**
Mills lambda	0.249	**
Observations	117	
Uncensored observations	95	
Wald chi²	37.25	

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

Table A.13a: Heckman two-step estimates[a] for social interactions, first most productive farmers dropped

Dep. Variable: Log(banana value)	(1) Kinship		(2) Neighbours <300m		(3) Social Insurance	
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Log(kibanja)	0.779	***	0.677	***	0.611	***
Log(adults)	0.761	x	0.521		0.870	*
Male grower	-0.082		-0.272		-0.323	
Some primary educ.	0.815	*	0.845	*	1.119	***
Completed primary	-0.001		-0.278		-0.248	
Log(age)	0.402		0.134		0.277	
Group averages:						
Log(productivity t-1)	0.330	(1.40)	-0.394		0.412	**
Constant	2.770		10.796	**	2.214	
Mills lambda	0.277		-0.339		0.548	
Observations	97		109		106	
Uncensored obs.	76		90		85	
Wald chi²	32.35		29.32		34.77	

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

[a] selection effects not shown in the table

Table A.13b: Heckman two-step estimates[a] for endogenous versus exogenous social interactions, first most productive farmers dropped

Dep. Variable: Log(banana value)	(1) Kinship		(2) Neighbours <300m		(3) Social Insurance	
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Log(kibanja)	0.847	***	0.585	***	0.604	***
Log(adults)	0.794	x	0.781	**	0.767	**
Male grower	-0.114		-0.155		-0.323	
Some primary educ.	0.934	**	0.918	**	1.057	**
Completed primary	-0.089		-0.376		-0.272	
Log(age)	0.408		0.109		0.187	
Group averages:						
Log(productivity t-1)	0.354	x	-0.814		0.369	
Head some primary	0.433		-1.446		0.706	
Head completed prim	-1.932	*	2.936		-0.757	
Log(age head)	-1.499		-2.982		0.071	
Constant	8.945		25.192	**	2.724	
Mills lambda	0.444		-0.232		0.330	
Observations	97		109		106	
Uncensored obs.	76		90		85	
Wald chi²	43.42		38.15		42.54	

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

[a] selection effects not shown in the table

Table A.14: Fixed effects regressions, two most productive farmers excluded

Dep: Log(banana)	(1) Kinship		(2) Neighbours <300m		(3) Social Insurance	
	Coeff.	Signif.	Coeff.	Signif.	Coeff.	Signif.
Log(adults)	-0.696	*	-0.584	*	-0.713	**
Log(prod t-1)	0.047		0.407		0.130	
Round3	0.388	***	0.342	***	0.397	***
Constant	7.862	***	4.580		7.161	***
Observations	99		155		134	
Nr. of groups	58		88		76	
F	5.65	***	9.86	***	7.39	***