Discussion Papers
Department of Economics
University of Copenhagen

No. 07-08

Social Interactions in Growing Bananas: Evidence from a Tanzanian Village

Katleen Van den Broeck
Stefan Dercon

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (online)

Social Interactions in Growing Bananas: Evidence from a Tanzanian Village

Katleen Van den Broeck (University of Copenhagen)

and

Stefan Dercon (University of Oxford)

April 2007

Abstract

This paper analyses whether agricultural information flows give rise to social learning effects in banana cultivation in Nyakatoke, a small Tanzanian village. Based on a village census, full information is available on socio-economic characteristics and banana production of farmer kinship members, neighbours and informal insurance group members. This allows a test for social learning within these groups and the identification of different types of social effects. Controlling for exogenous group characteristics, the effect of group behaviour on individual farmer output is studied. The results show that social effects are strongly dependent on the definition of the reference group. It emerges that no social effects are found in distance based groups, exogenous social effects linked to group education exist in informal insurance groups, and only kinship related groups generate the endogenous social effects that produce positive externalities in banana output.

JEL Classification: O12; O13; O55; Q12

Keywords: social interactions, social learning, agricultural information networks

Contact details: kvdb@econ.ku.dk or stefan.dercon@economics.ox.ac.uk.

1. Introduction

This paper investigates social interactions effects in the diffusion of pest-resistant agricultural technologies in a village in Kagera, Tanzania. To investigate technology diffusion, we have access to a unique full village level census of social networks, interactions and sources of knowledge. Furthermore, access to multiple observations on actual output over time allows us to focus not only on the adoption of the techniques, but also to quantify the impact of social effects on productivity. The village census allows us to explore a variety of possibly reference groups providing social effects.

In Kagera, in North-western Tanzania, population pressure is high and land is getting increasingly scarce. Soil fertility is decreasing and the local staple food, the Eastern African Highland banana, has become prone to diseases and pests. Farmer extension workers promote the use of technologies that mitigate the negative effects of banana diseases and weevil attacks. Although it appears that only few farmers are interested in learning techniques directly from the extension workers, social learning effects are likely to exist and produce positive externalities. In this paper we study whether and how interactions among farmers augment the effects of the few farmers directly interacting with extension officers.

A broad literature exists on the adoption of agricultural technologies (see e.g. Feder, Just and Zilberman (1985) and Besley and Case (1993)). Farmer characteristics such as human capital, degree of risk aversion and farm size are often found to influence adoption. But the assumption that farmer behaviour is also influenced by other farmers' actual behaviour or mere attitude towards a new technology has been gaining support. Although theoretical support of social effects in technology adoption has been quite common (for example, Ellison and Fudenberg, 1993, 1995; Gale, 1996; Bala and Goyal, 1998; Bardhan and Udry, 1999), empirical evidence measuring the importance of social effects is less extensive (Case, 1992; Besley and Case, 1994; Udry and Conley, 2001; Bandiera and Rasul, 2006; Munshi, 2004).

Especially sparse are the attempts to measure quantitatively the productive externalities of social learning. Where learning takes place, social effects are driven by information flows as opposed to mimicking or social pressure. As a way to distinguish the former from the latter Foster and Rosenzweig (1995) suggest analysing the effect of neighbours productivity enhancing behaviour on individual productivity rather than on individual (technique adoption) behaviour.

Another difficulty in empirical work is the definition of the relevant reference group. Not knowing the exact reference group seriously worsens the problem of identifying social effects as was argued by Manski (1993, 2000). Because of data constraints many authors use geographical boundaries to define farmer information networks. This may be the whole district as in Case (1992), the school as in Evans et al. (1992) or the village as in Foster and Rosenzweig (1995) or Munshi (2004). Also ethnicity has been used as reference group, for example by Borjas (1992, 1995). These groups are usually very large, but the relevant social interaction group may be much more limited.

Some findings suggest that farmer information reference groups are positioned at a smaller level than what has been studied before. Foster and Rosenzweig (1995) find for rural Indian farmers during the Green Revolution that villagers learn from neighbours experience in adopting high-yielding variety seeds but the effects are small compared to the effects of own experience. The authors suggest that villagers use the information of only a couple of neighbours.[1] Udry and Conley (2001) have also shown that information on the amount of inputs to use in pineapple cultivation in Ghana flows through networks not covering the whole village. Bandiera and Rasul (2006) show for sunflower adoption in Mozambique, that adoption decisions are correlated within family and friends groups but less so within religious based groups and not at all between groups of different religion. Inspired by these findings, this paper focuses on whether social effects in banana cultivation differ in type and magnitude according to the definition of the reference group.

A major innovation of the current study is that unlike most studies in the past, the Nyakatoke data were collected with the specific aim of analysing social effects.[2] The data take the form of a full village census of households and their interactions: all households in the village are included and precise information exists on different types of intra-village social networks. This allows us to use other than only

[1] 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 each agent is assumed to interact directly with only a finite number of other people.

geographical definitions of "the" reference group to investigate whether social learning effects exist and are of similar importance in different reference groups. The data allow for comparison among three social groups, namely kinship-based, distance-based and self-reported informal insurance groups.

Even when information on the reference group is available, identifying true social effects is not straightforward, as was shown in the seminal work of Manski (1993), and later by Brock and Durlauf (2001) and Durlauf (2002).[3] Especially where choice affects the formation of the group (endogenously formed groups), group effects have to be interpreted carefully as the results in Evans et al. (1992) show. Most of the groups considered do not suffer from endogenous group formation. Nevertheless, to distinguish endogenous and exogenous social effects, we use extensive exogenous characteristics of the group members.

This paper can improve considerably on earlier work. First, we have access to a full village census of households and their social networks. As a result, we do not need to rely on self-reported numbers of friends and family, such as whether they have adopted particular techniques, but can retrieve all this information from the interviews with all network partners. The village census of networks also allows us to distinguish possible learning effects in different networks, beyond the geographically defined effects usually studied. Secondly, the village census also allows us to test appropriately for social learning, by showing that not only a household's behaviour is affected by group behaviour, but also that its productivity is affected (Foster and Rosenzweig (1995)). The village census data include also lagged household level productivity data to implement our test of social learning. Furthermore, the village census allows us to identify the most productive farmers within their reference group, and which in turn can be excluded from the analysis to take into account the likely existence of a direction to social learning as suggested by Case (1992).

The results suggest that social effects in technique adoption exist in all groups but to a lesser extent in distance based reference groups. Analysing farmer banana output rather than technique adoption, it appears that exogenous social effects from group education exist in informal insurance groups, but the endogenous social effects which

---

[2] Data collection by author and Joachim De Weerdt (Katholieke Universiteit Leuven and Economic Development Initiatives, Bukoba, Tanzania) and a team of local enumerators.

[3] A survey article on social capital by Durlauf and Fafchamps (2004) treats identification and other issues that arise when analysing social capital in depth; a survey article by Soetevent (2006) focuses on the empirics of identification of social interactions.

produce a social multiplier effect, only exist in kinship related groups. The results are also suggestive of a direction to learning from the more to the less productive farmers in the group. These results suggest that social capital can play a positive role in household agricultural production parallel to its role in entrepreneurial activity (Barr, 2000; Fafchamps and Minten, 2002).

The remainder of the paper is structured as follows. Section 2 provides descriptive statistics of the learning and advice seeking behaviour of the farmers in Nyakatoke. Section 3 lays out the analytical framework and discusses the unique features of the dataset while section 4 presents the empirical specifications. Results are discussed in Section 5, and Section 6 concludes.

2. Information networks in Nyakatoke

Nyakatoke is a small village in the Kagera region of Tanzania, west of Lake Victoria. Banana is the main staple food and also an important cash crop for small scale farmers. The productivity of the indigenous banana trees has been declining for some years, mainly due to increased incidence of weevil attacks and panama disease. These pests are endemic, and have been gradually reducing the mean yield across all farmers in the village. Farmer extension officers are actively trying to introduce new and more resistant kinds of bananas, in combination with cultivation techniques that mitigate the effects of diseases. Although the output decline of the indigenous banana is perceived as a serious problem,[4] the reaction of farmers both in terms of adopting new types of bananas and in terms of using productivity enhancing techniques is still in an early stage.[5]

Data were collected during 2000 in five visits to the village with approximately two month intervals. Some data such as income and informal insurance assistance

[4] The productivity decline is seen as one of the major economic problems in the village. 26 per cent of farmers believe their output decreased slightly and 45 per cent believes it decreased strongly between 1990 and 2000. In Eastern Kagera (where the survey village is located) the average yield is only 3,100 kilograms per hectare, whereas average yields remain around 6,800 and 7,500 kilo per hectare in respectively Central and Western Kagera (where the soil is less depleted), see Agricultural Research Institute Maruku (1999).

[5] Only 22 out of the 119 households have adopted other banana types than the indigenous eastern African highland banana. Most households have planted them for experimental reasons and they usually have only one or two plants.

were collected in each round, while other data were only collected during one of the visits. No household sampling was necessary. The survey includes all 119 households and all income earning adults living in a household.

Two sets of data were collected to capture the sources of information on agricultural techniques. First, actual information flows were explored via questions on farmer knowledge, adoption and learning sources for 10 selected techniques, while secondly, hypothetical information flows were captured in the survey by a simple question "whom would you go to for advice if there was a problem with your crops?".

A list of 10 techniques for dealing with different banana problems was constructed based on the advice from extension workers.[6] Some of these technologies are straightforward to apply. Others are more time consuming or need specific knowledge, while still others require inputs such as manure. Some are complements, while others are substitutes. As the different pests affecting bananas are endemic, the main benefits of these techniques is to stem the decline in yields and not so much to reduce the variance in yields.

We found that on average, banana growing farmers knew between three and four of these techniques, although only 44 percent of the known technologies were actually used (or 1.6 techniques). Of all the households, 19 percent knew of none while 37 percent used none. There clearly is a large discrepancy between knowing and actually adopting a technique. The main reasons mentioned by respondents for not applying a certain technique were (1) doubts about the effectiveness of the technique, (2) having only recently learnt about it and not actually started applying it yet, (3) technique is too costly, or (4) difficulty of the technique.

To explore the sources of knowledge and adoption further, we pooled all the binary variables, so that each known technique is one observation. Table 1 shows the different learning sources and their relative importance. Some 57 per cent of the techniques known have been pulled into the village and 43 per cent have been learnt from another farmer or group of farmers in the village. A higher percentage of the known techniques is actually applied when farmers have learnt it from an outside source, likely to reflect self-selection. Farmers, who learn techniques from outside, are more likely to put some effort into looking for information and to undertake

[6] Techniques (re)introduced: special way of digging the hole (1), how to apply fertiliser/manure (2), hot water treatment of stem before planting (3), dipping stem in insecticide solution (4), how to mulch (5),

transaction costs to learn about the technique with a view to applying it. Farmers, who learn techniques from other farmers in the village, may do so accidentally, while visiting each other or discussing agricultural issues.

Table 1
Learning and diffusion of techniques[a]

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

[a] Several observations per household are possible, depending on the number of techniques known by the respondent (household head) so n=397 for the 119 households; [b] If the teacher of a technique could not be specified, it was either a dead person or a group of farmers.

For all techniques which were learned from farmers in the community, the name of the 'teacher' was obtained. This allowed us to construct Figure A1 in Annex which represents all farmers, who know at least one technique, and the information source for each technique they learned (so multiple sources per farmer can exist if techniques were learned from different sources). The outside sources include extension workers, relatives living in other villages (outside relatives) or friends living in other villages (outside farmers),[7] and other persons not known by name (outside no link). The existence of a two-layer structure can be observed. The first layer includes those farmers, who pull information into the village from outside. The second layer represents intra-village learning. It includes farmers, who receive their information from other community members. Most farmers appear to be linked directly or through only one other farmer to an outside learning source.

In the survey, we also collected detailed information on hypothetical advisers, based on the question "whom would you go to for advice if there was a problem with your crops?".[8] In Annex Figure A2 all farmers in the village are connected to their hypothetical adviser. The graph differs substantially from the graph in Annex Figure

---

trench-manuring (6), paring (7), desuckering (8), harvest hygiene (9), weevil trapping (10). See annex Table A1 for more detail on these techniques.

[7] In the case of "outside relatives", this sometimes includes deceased relatives such as parents or grandparents.

A1 where it concerns the number of layers. With respect to hypothetical advice seeking, there are more layers of intra-village links. Various farmers are multiple steps away from the outside source of advice. In any case, both figures confirm that behind farmers, who pull information into the village, there are numerous other farmers, who apply that knowledge.

In the remainder of this section, we explore this process further, and more specifically whether we can identify networks through which this information dissemination may take place. As part of the survey, detailed questions where asked to identify different social networks within the village. Here, we focus on three social groupings. First, a household's kinship network, defined as formed by all other households where at least one of the household members has a blood bond (up to the third degree) with the household in reference. Second, a household's informal insurance group includes all households on which the household reports that it can rely in times of need and vice versa. Third, a neighbourhood network, which is more arbitrary by definition, but consists here as households living within 300 metres of each other. The 300 metres threshold was chosen since it is lower than the average distance between households in the village (523 metres); a lower threshold would exclude many remote observations.[9]

It is of interest to reveal whether intra-village learning and advice links exist more among members of certain social groups than among random villagers. To do so, we consider all possible pairs (dyads) of household heads in the village, defining as one if a pair has a learning link (i.e. one learned a technique from the other), and zero otherwise. For hypothetical links, we construct a similar variable but consider all possible pair of adults involved in banana farming. The resulting link variable is explored in a dyadic regression framework (for example, as in De Weerdt (2004) or Fafchamps and Gubert (2007)), regressing the link variable on relational and non-relational variables (i.e. variables that describe the difference in a characteristic between a pair, and variables that control for the 'level' of the characteristic among the pair of adults). The latter type of variables is necessary to correct for type of

[8]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).

[9] To test sensitivity of the results, we chose the 10 closest households to serve as "the neighbourhood". All results were similar to those obtained by the 300 metres definition. Some coefficients were somewhat larger in magnitude, but they were not different in terms of significance, and not affecting the overall interpretation of our findings.

farmers we may be dealing with in the pair. For example, older farmers may have more experience in farming and therefore be more likely to be involved in information links. The same holds for larger farmers who may be better known or considered more innovative or knowledgeable. As relational variables we include the social group variables: whether the two persons belong to the same kin and to the same informal insurance network. Distance enters as a continuous variable measuring metres between the two homesteads. Additionally we include whether the two persons (individual respondents for hypothetical advice links, household heads for learning links) who are part of the pair are of the same sex, or age (if the difference is less or equal to 5 years), whether both have completed primary education or only one of both (base category is none of both), whether they live in households with the same land holdings (when the difference is less or equal to 0.5 hectares). For the advice link regression, which is at the individual level, we also include a dummy indicating that the two persons are living in the same household. The non-relational variables we include are the maximum age and the maximum landholdings of the farmers in the pair.[10] (we follow De Weerdt, 2004, in the use of maximum values to correct for non-relational effects).

Table 2 shows that being a member of the same kin or informal insurance network and a lower geographical distance between households positively and significantly affect the probability of an actual or hypothetical information link between two respondents. Further, farmers usually learn or go for advice to other farmers of the same sex. For actual learning links it turns out that the larger the farm of one of both farmers in the pair, the more likely there is a link. For hypothetical advice age/experience of the farmer has a positive effect. Also intrahousehold information flows appear to exist.

[10] We follow De Weerdt (2004) in the use of maximum values to correct for non-relational effects. Fafchamps and Gubert (2007) use the sum of the values of the pair. Using this correction for 'level' effects did not change the results.

Table 2
Information links marginal effects (probit regressions)

Dependent variable: 1 if link exists	Technique learning links (HH)[a]			Hypothetical advice links (IND)[b]		
	Coeff x 10^4	Std. err.[c]/signif.		Coeff x 10^4	Std. err.[c]/signif.	
Kinship related	42.448	22.998	***	6.399	3.444	***
Distance between HHs (metres)	-0.037	0.010	***	-0.007	0.003	***
Informal insurance members	35.147	18.692	***	53.025	13.103	***
Same age (difference <=5 years)	4.024	6.532		3.565	2.095	**
Same sex	8.598	4.701	*	7.526	1.719	***
Both completed primary	5.757	8.596		-1.378	1.240	
Only one completed primary	3.705	6.547		0.066	1.191	
Same land (difference <= 0.5 ha)[d]	-1.554	5.150		-1.437	1.147	
Maximum age	0.096	0.131		0.075	0.037	***
Maximum land (ha)	2.579	1.479	*	0.029	0.256	
Same household	-	-		7.367	6.752	*
Observations[e]	11342			43062		
Pseudo R^2	0.192			0.256		

[a]Technique learning links are analysed at household level; [b]Advice seeking links are analysed at individual level; [c]Robust standard errors, x 10^4; [d]Households have on average 1.22 hectares of land; [e]All households/individuals linked with all other households/individuals. A one directional learning link (link=1) exists for 40 combinations. Probability of an intra-village learning link: 0.3%. A one directional advice link exists for 86 combinations. Probability of an intra-village advice link: 0.2%.
*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 2 explored learning and advice links between farmers without taking into account the direction of the link, but as suggested by Case (1992) a direction may exist. This means that not everybody in the information network will be learning from everybody but some farmers will be teachers or advisers rather than “learners”. To explore this possibility we ranked all farmers in a network according to their agricultural productivity, assuming that whether a farmer is a teacher or a learner depends on his farming productivity. Table 3 shows indeed that farmers in the highest productivity quartile of the network are generally more likely to be mentioned as advisers or teachers than farmers in the lowest productivity quartile. This lends support for the existence of a learning or advising direction.

Table 3
Percentage teachers in lowest versus highest yield quartile farmers (within network)

	Kinship groups		Neighbours		Informal insurance groups	
	<25%	>75%	<25%	>75%	<25%	>75%
Mentioned as adviser	16	44	16	39	16	38
Mentioned as teacher	8	15	5	11	11	6

To conclude, the evidence suggests that information entering the village from an outside source, such as farmer extension services, will diffuse through the village, but rather than necessarily reaching all farmers in the village similarly (a common assumption in the literature where the village is often considered as the information reference group), it flows via networks.[11] The evidence for Nyakatoke suggests that social relationships such as kinship or informal insurance ties and physical distance are relevant proxies for farmer agricultural information networks. Furthermore, it appears that within a network the more highly productive farmers are more likely to be mentioned as teachers of techniques and as advice sources.

3. Formal framework

In this section a formal framework is presented showing how behaviour in a farmer social group can affect own behaviour through information flows and how this process can differ among groups. Problems of identification that arise when analysing social effects are also discussed.

3.1. Social effects in technique adoption behaviour

The theoretical framework on Bayesian updating of beliefs described by Berger (1985) and applied in the analysis of technology adoption for example by Besley and Case (1994), Foster and Rosenzweig (1995) and Udry and Conley (2001) forms the basis for the presentation of the empirical test of social interactions and learning.

Because there is only a certain percentage of farmers who actually apply a technique when knowing it, we assume farmer *i* initially believes that the production gain of applying the technique is not worth the effort. Each subsequent period reveals information about the true benefits of the technique through other farmers, who do apply it. Farmer *i* then updates prior beliefs using this information, and when farmer

[11] The assumption that the whole village constitutes one information network is further questioned through the following observation. Respondents were asked to indicate in a list of 10 persons whether they had actually gone for advice to this particular person the year before the survey and how many times. Persons included were the local extension officer, the innovator of the village (who was identified during focus group discussions), some of their informal insurance network members and one randomly drawn male and female farmer. Of all advice farmers obtained (total number of times farmers went to seek advice the year before the survey), the percentage that was obtained from these randomly drawn persons was close to zero. The percentage of advice obtained from a randomly drawn female farmer was 0, from a male farmer 4, from the local extension officer 8 and the rest of farmers’ advice came from the village innovator and the insurance members listed.

i's beliefs turn positive, the technique is adopted. If information is accurately transmitted farmer *i* is then also expected to obtain positive benefits from the technique. When this is the case, social learning effects exist.

In what follows, we will first present the case where information is transmitted without any noise and flows perfectly between all farmers in the village. Noise is subsequently introduced in the process, assuming that the variance of the noise is dependent on the relationship between farmers. The latter captures the observed effect that some relationships appear to enhance beneficial diffusion more than others.

The true benefit of a technique, X, identically and independently distributed across time and farmers, is assumed to be a normally distributed function, $X \sim \mathrm{N}\left(B, \sigma^2\right)$. All farmers know the variance σ^2 (no additional risk is associated with the techniques so farmers can infer the variance) but not the mean, B. They do hold beliefs about B. At *t-1* farmer *i* has prior beliefs on the benefits of applying a technique, B_{t-1}. Suppose the prior B_{t-1} is also normally distributed with mean β_{t-1} and variance τ_{t-1}^2. At time *t* farmer *i* updates existing beliefs about the benefits of the technique taking into account own prior beliefs and the information revealed by the experimenting farmers ($\overline{x}_t$ is the average benefit observed at time *t* of the n experimenting farmers). The posterior function $f(B|X)$ will be normally distributed with mean $\beta_t(x)$ and variance τ_t^2. If n farmers are experimenting and the information from their experiments reaches farmer *i* without noise, it follows from Bayesian updating rules that the updated expected benefit is:[12]

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

The process can only be determined if the parameters of the prior distribution, β_{t-1} and τ_{t-1}^2 are known. Reorganising (1) leads to:

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

Farmer *i* will update his beliefs downwards if the mean benefit obtained by the experimenting farmers is lower than the beliefs about the benefit held by farmer *i* and upwards in the other situation. When $\beta_t(x)$ becomes positive farmer *i* starts applying the technique.

The assumption of perfect and equal information has been shown to be rather unrealistic (Udry and Conley, 2001). (1) can therefore be modified to hold in a situation of imperfect information. Instead of observing x_t the farmer observes $x_t + u_t$. u_t is the measurement error, assumed to be independent from x_t and normally distributed with zero mean and variance equal to δ^2. 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 ($\mathrm{cov}(x,u) = 0$), $h(x,u)$ is $N(B, \sigma^2 + \delta^2)$. Assuming that the updating of beliefs is dependent on the relationship with the experimenting farmers, the variance of the noise is small where farmers have a close relationship, for example if they have known each other for a long time and trust each other. The variance of the noise is large where the information is revealed by farmers with whom farmer *i* has no special relationship. As such the information which is revealed by a group *k* of experimenting farmers is weighted differently, according to the relationship between the farmers in group *k*. Farmer prior beliefs about the benefits of a technique are updated in the following way:

$$\beta_t(x) = \beta_{t-1} + \frac{\tau_{t-1}^2}{\tau_{t-1}^2 + \dfrac{\sigma^2 + \delta_k^2}{n}} \left(\overline{x}_t - \beta_{t-1} \right) \tag{3}$$

where the coefficient on the updating factor is smaller due to noise and the updating speed is dependent on the relationship of farmers in the information network. In the limit, δ_k^2 goes to zero for a group *k* of highly trustworthy farmers and to infinity for farmers not trusted at all.

$$\text{For } \delta_k^2 \to 0: \ \beta_t(x) = \frac{\dfrac{\sigma^2}{n}}{\tau_{t-1}^2 + \dfrac{\sigma^2}{n}} \beta_{t-1} + \frac{\tau_{t-1}^2}{\tau_{t-1}^2 + \dfrac{\sigma^2}{n}} \overline{x}_t \tag{4}$$

[12] The updated variance is: $\tau_t^2 = \left(\frac{n}{\sigma^2} + \frac{1}{\tau_{t-1}^2} \right)^{-1}$

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

When δ_k^2 goes to infinity, farmer *i* refrains from using the information revealed by the experimenting farmer, that is prior beliefs are not updated. This theory lends support to the usefulness of exploring social effects across differently related groups.

3.2. Social learning test

Foster and Rosenzweig (1995) state that, in order to detect social learning it is necessary to show that an individual's productivity rather than behaviour is affected by reference group behaviour. As the final aim of the introduction of techniques is not technique diffusion per se but especially productivity increase, we test whether social learning in banana cultivation takes place. We model farmer output as follows:

$$Y_{i,t+1} = f(Z_{i,t}, \beta_{i,t}) \tag{6}$$

where $Y_{i,t+1}$ is the output of farmer *i* in period *t+1*. $Z_{i,t}$ is a vector of characteristics of the farmer and of the land suitable to cultivate bananas. $\beta_{i,t}$ is farmer *i*'s prior belief about the benefit (or expected benefit) of applying a technique. It contains a history of private information in combination with information on experiences of other farmers and determines whether and how a technique is applied. As shown in (3) $\beta_{i,t}$ is a combination of the beliefs of farmer *i* prior to experimentation, of the average benefit on experimenting farmers' fields and on the relation between farmer *i* and the experimenting farmers. For the test, we have to make an assumption on $\beta_{i,t-1}$. We assume that farmers own prior beliefs about the benefit of a technique are neutral. Therefore we only include the average benefit on the fields of *n* information network members ($\overline{X}_{-i,t}$), corrected for the relevant weighting factor $\dfrac{\tau_{t-1}^2}{\tau_{t-1}^2 + \dfrac{\sigma^2}{n} + \dfrac{\delta_k^2}{n}}$ in group *k*. Thus we will test for different networks *k*:

$$Y_{i,t+1} = f(Z_{i,t}, \overline{X}_{-i,t,k}) \tag{7}$$

To bring this to the data, we will use a log-linear form, resulting in the following equation of interest:

$$\log(Y_{i,t+1}) = \alpha_0 + \alpha_1 \log(Z_{i,t}) + \alpha_{2,k} \log(\overline{X}_{-i,t,k}) + u_{i,t} \quad \text{for } k=1,\ldots,K \tag{8}$$

If farmers in group *k* learn from each other, $\log(\overline{X}_{-i,t,k})$ is expected to have a positive effect on $\log(Y_{i,t+1})$ and to have no significant effect otherwise.

3.3. Identification of social effects

The identification of endogenous social effects such as the effect specified in (8) is not straightforward. There has been much debate on exactly how society affects an individual and whether "real" social effects in economic decision-making actually exist. [13] The seminal work by Manski (1993, 2000) tackles the identification problem theoretically. He describes and formalizes three hypotheses to explain the observation that individuals in the same social group tend to behave similarly, only one of which is endogenous and produces the so-called social multiplier effects. First, endogenous social effects cause individual behaviour to vary with group behaviour. There is an endogenous effect if, ceteris paribus, individual outcome tends to vary with group achievement. Second, there can be exogenous social effects, where individual behaviour varies with the exogenous group characteristics. Third, correlated effects exist where individuals in the same group behave similarly because they have similar individual characteristics or face the same environment. The last effect is not a social effect. The 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. In this case the programme does not only affect the achievement of the tutored

[13] Social effects are not only analysed in adoption behaviour or in output performance, where the channel of social interactions is mainly information, but also in other types of behaviour where the social interaction may rather be exhibited via norms. For example, Borjas (1992) tested the effect of average earnings within the same ethnic group as the parents on current earnings of the children, which was assumed to work via peer group pressure on the parents. Bertrand, e.a. (2000) used a similar basic test to analyse welfare benefit use and how being part of a social group may inhibit mobility. Krishnan (2001) analysed the fertility behaviour of Indian women. Yamauchi (2007) shows that learning about returns to schooling is faster when the variation in observed characteristics is larger.

students, but indirectly also affects the achievement of other students in the group. The other two effects do not generate this social multiplier. In the case of technique promotion, only the existence of endogenous social effects will augment the effort of teaching techniques to a limited number of interested farmers by increasing not only the output of the contacted farmers but also the output of other farmers in the contact farmers reference groups.

In Manski (1993, 2000) there are three important econometric issues to take into account when trying to identify endogenous social effects: (1) the simultaneity or reflection problem, (2) the problem of identifying the exact reference group and (3) the omitted variables problem. A problem is how to model social effects empirically taking these issues into account as discussed by Brock and Durlauf (2001). In what follows we discuss the three issues in turn and we will highlight how the specific features of the Nyakatoke dataset can be used to approach the problems.

First, the reflection problem arises because the behaviour of farmers in the reference group affects the behaviour of an individual farmer in that group, but the behaviour of that farmer in turn affects group behaviour, causing simultaneity. One of the possible solutions Manski offers is to make the model dynamic and assume a lag in the transmission of social effects.[14] Including lagged group behaviour instead of contemporaneous group behaviour is necessary but not sufficient for identification. Using lagged group behaviour only offers a solution to the identification problem when the process of social effects is observed out of equilibrium. Moreover, the timing of the lag has to be established. Fortunately, both these conditions are satisfied in the Nyakatoke data set. Only few farmers apply the techniques, and farmers generally feel they are not necessary yet or too difficult to apply. This suggests an out-of-equilibrium situation of banana cultivation methods. Furthermore, data were collected on current banana harvests but also retroactively for the previous year. This one year time lag is appropriate since the banana plants take approximately one year to become fully grown and flowering and before the results of technique use may be visible in the harvested banana bunches.

[14] Another alternative Manski and others propose is to use a non-linear model, which 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.

Second, identifying social effects strongly depends on the identification of the reference group. In most of the existing empirical literature farmers living in the same neighbourhood, the same village, district or ethnic group are assumed to be group members. However, these are all *expected* interaction groups. If the true information networks do not match these expectations, the village averages used in empirical analysis are only estimates of individual specific variables. Moreover, they may not show much variation for observations within the same geographical boundaries. Mostly these expected information groups are exclusive: an individual belongs to only one interaction group. Recently, attempts have been made to collect information on the exact interaction groups which farmers belong to. Udry and Conley (2000) collected data on pineapple growing farmers in Ghana. They asked whether and how many times the interviewed farmers had talked to a list of persons. Their data show that information networks do not necessarily cover a whole village or are mutually exclusive, but may show complicated patterns. The Nyakatoke data contain individual level information on actual social groups farmers are part of (kinship networks, geographical networks and self-reported informal insurance networks) and these characteristics were shown in Section 2 to be reasonable proxies for farmer information networks. Therefore they are used as individual level reference groups in the empirical analysis in Section 5.

Third, the estimation of social effects often suffers from an omitted variables bias, which leads to biased estimates of the coefficient on $\overline{X}_{-i,t,k}$. Equation (8) represents the pure endogenous effects model where it is assumed that exogenous effects are non-existent. This assumption is a very strong one, and is typically made because data on exogenous characteristics of group members are not readily available. Where social groups are endogenously formed, the problem is even more serious. So far we have assumed that the rules of group formation do not have any effect on the identification of social effects, but when possible, individuals will endogenously sort themselves into groups. For example, farmers will try to link up with farmers, who have certain characteristics or abilities, and De Weerdt (2004) finds for Nyakatoke that characteristics such as kinship, clan, distance, education and wealth correlate with informal insurance group formation. When endogenous matching takes place, there is strong potential for self-selection bias,[15] and the endogenous social effects may be misinterpreted. The variables that drive group formation may also drive farmers' outcomes, but there are no endogenous social effects present. Since we have information on exogenous characteristics of *all* intra-village group members, we can explicitly include the exogenous group characteristics as controls ($Z_{-i,t,k}$) together with average group behaviour ($\overline{X}_{-i,t,k}$) and the individual controls ($Z_{i,t}$), so the social learning test becomes:

$$\log(Y_{i,t+1}) = \alpha_0 + \alpha_1 \log(Z_{i,t}) + \alpha_{2,k} \log(\overline{X}_{-i,t,k}) + \alpha_{3,k} \log(Z_{-i,t,k}) + u_{i,t} \tag{9}$$

Identification of the different social effects is then ensured if there is no multicollinearity present. It holds that: (1) If $\alpha_1 \neq 0$ correlated effects exist; (2) if $\alpha_{2,k} \neq 0$ there are endogenous social effects in group *k*; and (3) if $\alpha_{3,k} \neq 0$, exogenous social effects exist in group *k*. Only in case (2) will farmer behaviour vary with behaviour in the reference group. Only this effect produces social multiplying effects in banana cultivation.

The advantage of working with data on exact group composition and lagged average group behaviour is rare in the empirical social interactions literature, and to our knowledge it is unique to have information on different types of individual specific reference groups and the exogenous characteristics of each of the members.[16]

4. Empirical model

The existence and type of social interactions will be tested for two specifications: technique adoption behaviour and banana output. First, we test whether the technique adoption behaviour of a farmer (A_i) is influenced by the technique adoption behaviour in the reference group (A_{-i}):

$$A_i = \alpha_0 + \alpha_1 Z_i + \alpha_{2,k} A_{-i} + \alpha_{3,k} Z_{-i} + u_i \tag{10}$$

[15] 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.
[16] Although the data have many advantages for analysing social effects, they also have shortcomings such as being a small and geographically very concentrated sample.

The dependent variable is one if the household uses the technique and zero otherwise. The reference group behaviour A_{-i} is captured by the *number of group members using technique*. Unfortunately, no information was collected on the year of adoption, so we are forced to study group and individual behaviour contemporaneously. Z_i is a vector of individual characteristics, Z_{-i} is a vector of exogenous characteristics of farmer *i*'s reference group (farmer *i* is always excluded from variables that capture group behaviour or group characteristics) and u_i is $N(0,\sigma_u^2)$. The endogeneity and reflection problem is in this case potentially serious, as controlling for all possible exogenous variables that may determine both group and individual adoption may not be sufficiently credible, not least if individuals self-select into groups, so that individuals with a shared interest in innovation may be part of the same group. However, the nature of the groups used makes this less likely to be a problem, definitely for two of the groupings chosen as membership is exogenously determined: kinship, and given traditional communal land allocation, neighbourhood. The problem may be more important for the informal insurance network, even though its purpose of existence appears very different from adoption of technologies. Overall, however, we cannot claim too much from the regressions based on (10).

Vector Z_i contains household level information on (1) land suited for banana cultivation[17] (*kibanja*), (2) the quality of the land, measured by two dummy variables capturing whether land quality is higher than the average in the village (*high quality*) or average (*average quality*) as reported by the household (3) the number of *adults* present to capture labour availability, and (4) a dummy for radio ownership (*radio*) representing the household's access to information. Individual characteristics are (1) gender (*sex*) to correct for possible gender biases in banana cultivation, and (2) *age*, to capture experience in banana growing.[18] Furthermore, two dummies are included to

[17] This type of plots is called the kibanja in Swahili. See ARI Maruku (1999, p.45), Mitti and Rweyemamu (2001, p.15), and Maruo (2002, p.151), for characteristics of the kibanja and other plot types and the type of crops grown on different plot types. The number of hectares that can be cultivated with banana trees is usually restricted to kibanja, which is the plot around the house. 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 high population density, and there is not much opportunity to expand the kibanja. The kibanja is also cultivated with coffee, maize and beans, all intercropped. There is no information on the cultivated area separately for each household member, so total household kibanja land is used. Individual information would be nearly impossible to collect since husband and wife often care for all the banana trees and other crops in the kibanja together.

[18] In (10) individual level characteristics are the characteristics of the household head but in (11) they are the characteristics of the banana grower.

capture whether the farmer completed (1) *lower primary schooling* (first four years) and (2) *higher primary schooling* (years five to seven).

Vector Z_{-i} (exogenous characteristics of the reference group) includes (1) the number of household heads in the reference group, who completed lower primary schooling (*number of heads lower primary*), (2) the number of heads in the reference group, who have completed higher primary schooling (*number of heads higher primary*), (3) the average age of the household heads in the reference group (*average age heads*), (4) the *number of male headed* households in the reference group, (5) group members average size of land suitable for growing bananas (*average kibanja size*), and to correct for group size, (6) the number of members in the whole group (*group size*) and (7) its squared term (*group size*^{2}) are also included.

Finding social effects in (10) may well be the result of mimicking or social pressure. To test whether social learning exists in a credible way, we test whether individual actual outcomes in period *t* ($Y_{i,t}$) are affected by the average benefits of technique use in the reference group in the *previous* period by using a modification of the log-linearised Cobb-Douglas production function in (9):

$$\log(Y_{i,t}) = \alpha_0 + \alpha_1 \log(Z_i) + \alpha_{2,k} \log(\overline{P}_{-i,t-1}) + \alpha_{3,k} \log(Z_{-i}) + u_{i,t} \tag{11}$$

where Z_i and Z_{-i} are as previously described. $\overline{P}_{-i,t-1}$ is the *average banana yield* (banana output per hectare) at time *t-1* of the households that form part of farmer *i*'s reference group. It serves as a proxy for the actual average group benefit of technique adoption. More correctly, it captures the output per hectare effect of the total of agricultural methods applied by group members. We use the average yield of the households that are linked to the household farmer *i* is part of, so intra-household information pooling is assumed (and suggested by the results in Table 2). If any household member receives information on banana cultivation, this information will be shared with fellow banana growing household members.[19] We use average yields, and not other moments of the distribution of yields as the main impact of the improved techniques is to stem the decline in yields, and not so much the variance, as

[19] Prices are not included. The price of a bunch of bananas does not vary between neighbourhoods in the village and is roughly equal throughout the year. There is no specific time for harvesting bananas, which are grown continuously during the year. Although the banana output is captured in monetary values, differences truly reflect quantitative differences.

was discussed in section 2. Using lagged values allows us to more convincingly capture social effects going from the reference group onto the farmer.

Table 4 contains selected descriptive statistics of individual characteristics of banana growers and the number of group members for reference group *k.* The reference groups are as defined previously: kinship related farmers, neighbours and informal insurance members. Although bananas are grown by all households, only 95 persons (81% of all banana growers) mentioned selling bananas as an income earning activity. For banana selling respondents we have information on total output values, including both the amount sold and the amount consumed by the household but for banana growers who do not sell any surplus we have neither so these observations will drop from the analysis.[20] Table 4 also shows that some banana growers do not have members of the same kin in the village which will further reduce the number of observations due to missing variables for group characteristics.

Table 4
Characteristics of banana growers

Variables	Average	Observations
Individual characteristics of 117 banana growers:		
Male growers (%)	50	117
Age (years)	43	117
Lower primary education (% completed standard 1 to 4)	77	117
Higher primary education (% completed standard 5 to 7)	56	117
Banana sellers (as % of banana growers)	81	117
Total banana output value of sellers (in Tsh[a])	6,841	95
Of all households with at least one banana grower:		
Kibanja (hectare)	0.5	101
Total land holdings (hectare)	1.2	101
Adults present in the household (> 15 years)	2	101
Radio ownership (%)	36	101
Social groups of banana sellers (nr of persons in group):		
Number of kinship members	8	88[b]
Banana selling kinship members 1999	5	88
Number of neighbours (living within less than 300 metres)	32	95
Banana selling neighbours 1999	21	95
Number of informal insurance network members	10	95
Banana selling informal insurance network members 1999	7	95

[a] Tanzanian Shilling (1 US$=+/-800Tsh; in 2000); [b] There are some banana sellers who do not have any kin related households in the village, hence the lower number of observations.

[20] Due to the design of the survey we do not have information on output values for those households who did not mention banana cultivation as a source of cash income, as the survey aimed to capture only the income earning activities in which individuals were engaged in, in a context of a high diversity of income earning activities. This observation suggests a possible sample selection problem (missing output data for non-selling banana growers), although no obvious candidates for identifying instruments exist in the data so investigating this further proved not feasible.

5. Estimation results

The results of the specifications introduced above will be discussed in turn. Tables 5 and 6 show the effects of group variables on individual behaviour and output while the effects of individual farmer and land characteristics are shown in the Annex (Tables A2 and A3). To explore technique adoption behaviour we run a probit estimation on pooled techniques correcting for household clustering.[21] With respect to individual level effects on household adoption of techniques (see Annex Table A2) we find that completing higher primary education and ownership of a radio are consistently positive determinants of technique adoption. Additional to individual characteristics, group level effects also appear to play a role in individual adoption behaviour (shown in Table 5).

Although the coefficients differ in magnitude, the number of reference group members using a technique has a positive and significant effect for all types of reference groups (after correction for group size). An additional adopting kin or informal insurance network member increases the probability of a farmer adopting a technique by 6 per cent. Distance based groups show a lower impact of only 2 per cent. A similar result was found by Isham (2002) analysing fertiliser adoption of Tanzanian farmers. It appeared that households with ethnically based affiliations were more likely to diffuse the technology successfully. Exogenous social effects are found only in informal insurance networks. An additional household in the group with a head, who completed lower primary school, positively affects the individual probability of adoption by 3 per cent. Although correlations consistent with social effects appear to exist in technique adoption, it needs to be established whether this translates into social effects in banana outcomes.

[21] We chose to pool techniques rather than analysing the number of techniques adopted via ordered probit. The reason is that it is not clear whether more techniques are better than only a few. The techniques are suited to mitigate the effects of a range of different banana growing problems such as the black sigatoka disease or weevil attacks. Dependent on the type of problem a farmer faces, a certain technique will be beneficial.

Table 5
Marginal group effects on technique adoption, pooled techniques[a]

Dependent variable: use of technique 0/1	Kinship	Neighbours <300m	Informal insurance network
Group characteristics			
Number of group members using technique (-*i*)	0.064***	0.022***	0.059***
	(0.009)	(0.003)	(0.009)
Average age heads	-0.000	0.003	-0.003
	(0.002)	(0.006)	(0.003)
Number of male heads	-0.014	-0.005	-0.027*
	(0.019)	(0.015)	(0.015)
Heads with lower primary	0.007	-0.004	0.030*
	(0.025)	(0.011)	(0.016)
Heads with higher primary	-0.010	0.004	0.015
	(0.017)	(0.008)	(0.016)
Average kibanja size in group	-0.074	0.081	-0.038
	(0.071)	(0.207)	(0.090)
Group size	-0.003	-0.005	-0.016
	(0.021)	(0.008)	(0.018)
(Group size)^2	-0.000	0.000**	-0.000
	(0.001)	(0.000)	(0.000)
Observations	1003	1003	1003
Pseudo R^2	0.16	0.22	0.22

Robust standard errors in parentheses; adjusted for household clustering. Effects of household characteristics are shown in Annex Table A2.
* significant at 10%; ** significant at 5%; *** significant at 1%.
[a] Probit marginal effects. There are 10 observations per household since technique adoption was asked for a series of 10 techniques. Number of observations limited by banana growing households with banana growing kin members (101); further limited by 7 missing observations on technique use.

To test social interactions in banana output using specification (11) all continuous variables are in logarithms, and dummies remain in levels. Variables with a lot of zero observations remain in the regression as discrete variables to prevent losing many observations while taking logarithms.

In Table 6 we present the results of the social interactions test, which allows individual output to vary with average group yield and exogenous group and individual characteristics. Assuming that farmers prefer to learn from their high yield group members rather than from poorer performing farmers (as suggested by the data, see Table 3), as part of our test, the highest yield farmer in each reference group is excluded from the analysis all together. Including “teachers” may underestimate the social effects.[22] This is tested by comparing the results when all farmers are included (columns 1 to 3) with the results when the highest productive farmer in a group is

[22] Farmers with the highest agricultural productivity in a group are more likely to teach other group members than to learn from them. Including these observations disregards that a direction to learning may exist.

excluded (columns 4 to 6). Individual effects are presented in Table A3 and highlight the importance of land size, land quality and completing lower primary education.

There is evidence of social effects but again it differs by group. The single reference group for which endogenous social effects clearly exist is based on kinship links. In informal insurance groups individual behaviour varies only with exogenous characteristics of the group, namely the number of members who completed lower primary education and the average kibanja size of the group members. In distance based groups no social learning effects seem to be present.[23]

Table 6 lends support to the hypothesis that there is a direction to the social effects going from the better to the less performing group members. Both the magnitude and the significance of the endogenous social effect found in kinship groups increase when best performing farmers are excluded from the analysis. The exogenous social effects of lower primary education in informal insurance networks also increase in magnitude when high productive farmers are excluded.

The endogenous social effect found in kinship based groups suggests there is potential for positive externalities of efforts to increase productivity of only a few well-targeted farmers. In informal insurance groups it is the stock of education in the group that positively influences individual member output. This effect is comparable to the positive intra-household externality of a literate household member described by Basu and Foster (1998). In the distance based group where no social effects are present, farmers appear to rely more on their own education (see Table A3). In both kinship related and informal insurance related groups, having larger farm holders as members positively affects individual outcome. This may also capture an information externality where farm size is associated with a higher level of innovation and technique adoption.

[23] All group results hold, even when lagged own productivity is included, which is itself significant.

Table 6
Effect of average group yield on individual banana outcomes[a]

Dependent: Log(banana harvest value)	All farmers in network included			Highest productive farmer in network excluded		
	Kinship	Neighbours <300m	Informal insurance network	Kinship	Neighbours <300m	Informal insurance network
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Average banana yield *t-1*)	0.265	-0.571	0.054	0.466**	-0.33	0.108
	(0.191)	(0.431)	(0.170)	(0.225)	(0.576)	(0.195)
Log(Average age)	-1.294	-1.584	-0.071	-1.788	-2.335	-0.013
	(1.318)	(2.732)	(0.994)	(1.418)	(2.963)	(1.022)
Number male heads	-0.432***	-0.128	-0.178	-0.476***	-0.118	-0.127
	(0.164)	(0.168)	(0.132)	(0.178)	(0.173)	(0.140)
Heads lower primary	0.202	0.218	0.301**	0.195	0.209	0.326**
	(0.262)	(0.132)	(0.146)	(0.295)	(0.137)	(0.153)
Heads higher primary	-0.201	-0.063	-0.090	-0.203	-0.071	-0.180
	(0.172)	(0.074)	(0.124)	(0.196)	(0.079)	(0.133)
Log(Avg kibanja ha)	0.844	0.983	1.199**	1.333**	1.183	1.201**
	(0.560)	(1.415)	(0.472)	(0.597)	(1.479)	(0.501)
Group size	0.241	0.002	-0.134	0.270	0.016	-0.165
	(0.173)	(0.100)	(0.152)	(0.181)	(0.105)	(0.157)
$(\text{Group size})^2$	0.002	-0.000	0.005**	0.002	-0.000	0.006**
	(0.007)	(0.001)	(0.003)	(0.008)	(0.001)	(0.003)
Constant	9.872	15.561	7.428*	10.980	16.714	8.168*
	(6.234)	(12.109)	(4.323)	(6.575)	(14.092)	(4.408)
Observations	87	95	95	76	90	85
Adjusted R-squared	0.36	0.32	0.39	0.34	0.29	0.36

Standard errors in parentheses. Effects of individual characteristics can be found in Annex Table A3.
* significant at 10%; ** significant at 5%; *** significant at 1%

The table supports the idea that social capital is truly capital in the sense that it contributes to production as suggested in Narayan and Pritchett (1999). The empirical literature on the productive effects of social capital is generally limited to entrepreneurial activity. Unlike the result found in Fafchamps and Minten (2002) where family based network relationships appear to reduce firm productivity amongst traders in Madagascar, in agricultural households in Nyakatoke family based relationships seem to have more beneficial effects. But even though the results obtained here are suggestive of the existence of social effects, it has to be borne in mind that they may look different or may even be non-existent when analysing other villages, other types of reference groups or other crops. For example Munshi (2004) showed that social learning effects are crop specific and are different for wheat and rice growers in India.

A remaining question is why the endogenous social effect only exists in kinship related groups. The answer possibly lies in the fact that, in order to gain from the knowledge of other farmers, knowledge has to be passed on very meticulously. Farmers need to know the exact way in which good results are obtained.[24] Presumably kinship related farmers, such as parents and siblings put more effort into explaining the technology than would neighbours or informal insurance network members do. Moreover, among kin-related farmers there may be fewer unobserved characteristics which can reduce information flows.[25]

6. Conclusion

This paper departed from the observation of an apparently slow diffusion process of productivity enhancing techniques in banana cultivation in Nyakatoke, a village in north-western Tanzania. This may cast doubt on the effectiveness of extension activities. We have highlighted the possibility that there are positive externalities of the few farmers who visit extension centres and learn new banana cultivation methods. Moreover, we studied whether these externalities are dependent on the relationship between farmers.

The analytical framework involved Bayesian updating of beliefs with regard to the benefit of technique adoption. The updating process is made dependent on the relationship between farmers. A social interaction test of the Manski type was derived where individual output varies with (1) average group achievement (endogenous social effects), (2) group characteristics (exogenous social effects), and (3) individual characteristics (correlated effects). Due to the nature of our dataset, solutions were found to tackle potential econometric problems which usually hamper the identification of different types of social effects. For example, lagged group achievement instead of contemporaneous achievement was used to address simultaneity. Moreover, since the data were collected with the specific aim of analysing social groups the exact composition of three different reference groups is known and data on the achievement and socio-economic characteristics of all

[24] For example, a technique exists to prevent weevils from attacking the banana plant. 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 trap at regular times, e.g. at least once a day. But some periods of the day are known to be better suited to clean these weevil traps.

reference group members is available. This enabled us to use individual specific reference groups and include exogenous group controls to tackle standard difficulties related to identification and omitted variables bias respectively.

Rather than studying one social group as is usually the case, three different reference groups were tested for the existence of social effects. In order of exogeneity these are (1) kinship related groups, (2) neighbourhoods and (3) self-reported informal insurance groups. Social effects appeared to exist in some, but not all groups, and the nature of the effects was quite different according to the group under analysis. Household technique adoption behaviour was strongly affected by the number of adopting group members in all groups, but the group effect was smaller in neighbourhoods. Finding social effects in technique adoption behaviour did not appear to automatically give rise to social effects in banana output. Endogenous social effects only existed within kinship related reference groups. Although no endogenous social effects were found within informal insurance networks, there are exogenous social effects of education and group members average banana land size. All social effects appear to work from the better to the less performing group members.

In sum, the results highlight especially that the definition of the reference group plays an important role in the identification of social effects. For the survey village of Nyakatoke in north-western Tanzania the results suggest that social effects in technique adoption exist in all reference groups tested but information transmission only has social multiplier effects on farmer's output in one of the reference groups. To obtain the widest output externalities of teaching techniques to a limited group of farmers, it might be necessary to choose as contact farmers those belonging to different kinship groups within which information diffusion has better output results.

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

Acknowledgements

We are thankful for the useful comments received on early or later drafts of the paper from Thomas Barnebeck-Andersen, Mikkel Barslund, Lode Berlage, Joachim De Weerdt, Geert Dhaene, Frédéric Gaspart, Jo Swinnen, Finn Tarp, and Bjorn Van Campenhout. We benefited from the suggestions of seminar participants at CES-Leuven, IFPRI, and CAM in Copenhagen. This work would not have been possible without the participation and hospitality of the people of Nyakatoke so we would like to express our gratitude to them and to the team of enumerators of the Nyakatoke Household Survey. A grant from the Flemish Interuniversity Council for data collection is gratefully acknowledged. All errors are our own.

References

Agricultural Research Institute Maruku, 1999. Planning the Future: Past, Present and Future Perspectives of Land Use in the Kagera Region. ARI Maruku, Bukoba and Royal Tropical Institute, Amsterdam.

Bala, V., Goyal, S., 1998. Learning from Neighbours. The Review of Economic Studies 65, 595-621.

Bandiera, O., Rasul, I., 2006. Social Networks and Technology Adoption in Northern Mozambique. The Economic Journal 116(514), 869-902.

Bardhan, P., Udry, C., 1999. Development Microeconomics. Oxford University Press.

Barr, A., 2000. Social Capital and Technical Information Flows in the Ghanaian Manufacturing Sector. Oxford Economic Papers 52, 539-559.

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

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

Besley, T., Case, A., 1993. Modelling Technology Adoption in Developing Countries. American Economic Review 83(2), 396-402.

Besley, T., Case, A., 1994. Diffusion as a Learning Process: Evidence from HYV Cotton. RPDS Discussion paper 174, Princeton University, New Jersey.

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

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

Brock, W.A., Durlauf, S.N., 2001. Interactions-based models. In: Heckman, J., Leamer, E. (Eds), Handbook of Econometrics, volume 5. Elsevier, Amsterdam, 3297-3380

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

De Weerdt, J., 2004. Endogenous Network Formation, chapter in S.Dercon (ed.), Insurance against Poverty, Oxford University Press.

Durlauf, S.N., 2002. On the Empirics of Social Capital. The Economic Journal 112, F459-F479.

Durlauf, S.N., Fafchamps, M., 2004. Social Capital. National Bureau of Economic Research, NBER Working Papers : 10485.

Ellison, G., Fudenberg, D., 1993. Rules of Thumb for Social Learning. Journal of Political Economy 101(4), 612-643.

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

Evans, W.N., Oates, W.E., Schwab, R.M., 1992. Measuring Peer Group Effects: A Study of Teenage Behavior. Journal of Political Economy 100(5), 966-991.

Fafchamps, M., Minten, B., 2002. Returns to Social Network Capital among Traders. Oxford Economic Papers 54, 173-206.

Fafchamps, M., Gubert, F., 2007, The Formation of Risk-Sharing Networks, Journal of Development Economics 83(2), 326-350.

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

Foster, J.E., Basu, K., 1998. On Measuring Literacy. Economic Journal 108, 1733-1749.

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

Gale, D., 1996. What Have we Learned from Social Learning? European Economic Review 40, 617-628.

Isham, J., 2002. The Effect of Social Capital on Fertiliser Adoption: Evidence from Rural Tanzania. Journal of African Economies 11(1), 39-60.

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 60, 531-542.

Manski, C., 2000. Economic Analysis of Social Interactions. Journal of Economic Perspectives 14(3), 115-136.

Maruo, S., 2002. Differentiation of Subsistence Farming Patterns Among the Haya Banana Growers in Northwestern Tanzania. African Study Monographs 23(4), 147-175.

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

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

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

Rogers, E.M., 1995. Diffusion of Innovations. The Free Press, New York.

Soetevent, A.R., 2006. Empirics of the Identification of Social Interactions: An Evaluation of the Approaches and Their Results. Journal of Economic Surveys 20(2), 193-228.

Udry, C., Conley, T., 2000. Learning About a New Technology: Pineapple in Ghana. Economic Growth Center, Discussion Paper: 817, Yale University.

Udry, C., Conley, T., 2001. Social Learning through Networks: The Adoption of New Agricultural Technologies in Ghana. American Journal of Agricultural Economics 83(3), 668-673.

Yamauchi, F., 2007. Social Learning, Neighborhood Effects, and Investment in Human Capital: Evidence from Green-Revolution India. Journal of Development Economics 83(1), 37-62.

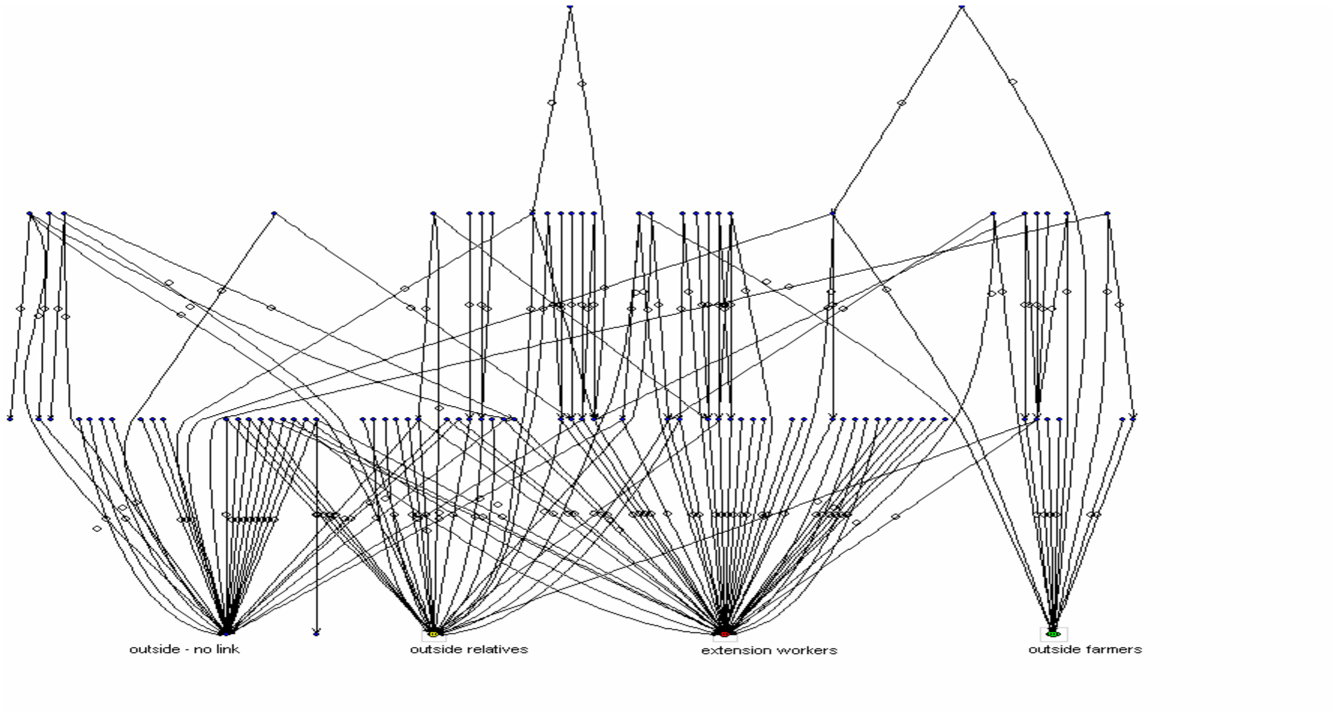
Annexes

Table A1
Explanation of techniques to maintain banana plants

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.

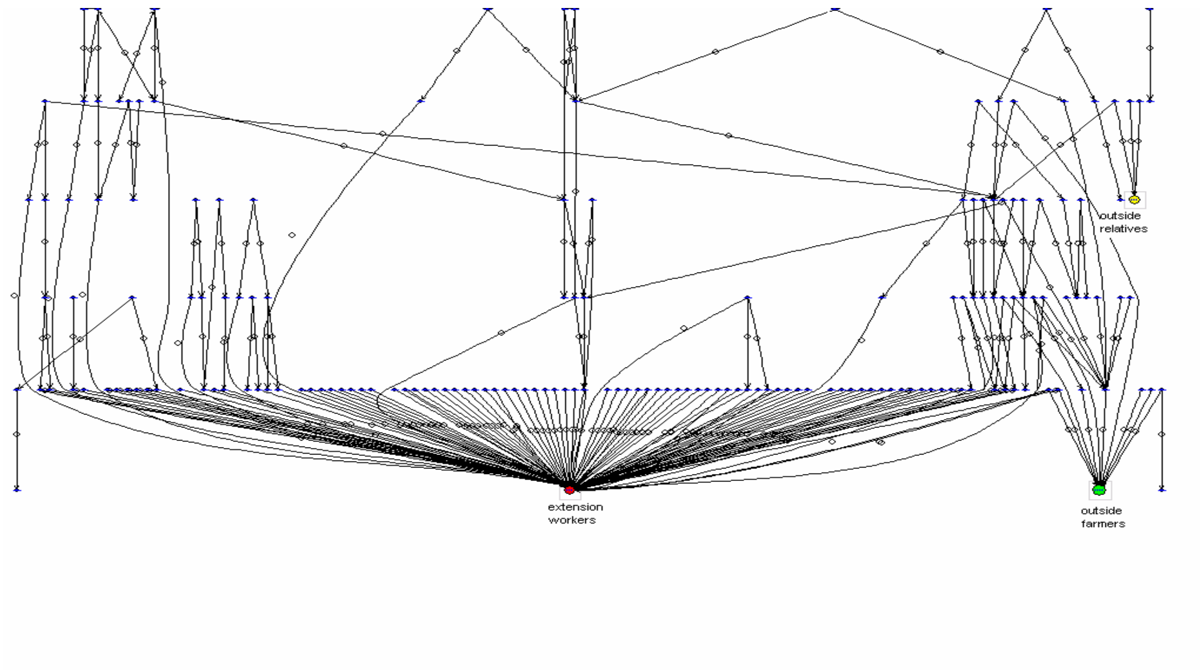
Source: Mbwana, A.S.S e.a. (1998), “A Guide to Growing Bananas in the Eastern African Highlands”, ICIPE

Figure A1
Technique diffusion



Note: Data on 397 techniques known and the source of information. Farmers (small black dots) are connected to their technique learning source, i.e. other farmers in the village (known by name), farmer extension workers, relatives not living in the village, other farmers not living in the village but known by name and other persons not living in the village and not known by name. The bottom layer includes out-of-village sources while upper layer black dots are all farmers interviewed who know at least one technique.

Figure A2
Hypothetical advice links



Note: Farmers are connected to their advice source. Outside advice sources are extension workers or other farmers. Other dots are farmers interviewed as organised by software programme "dotty".

Table A2
Household technique adoption: farmer effects[a] – to Table 5.

Dependent variable: use of technique (0/1)	Kinship	Neighbours <300m	Informal insurance network
Household characteristics			
Kibanja area (hectare)	0.093**	0.072	0.039
	(0.040)	(0.044)	(0.042)
Land perceived of high quality	-0.009	0.085	0.010
	(0.041)	(0.077)	(0.047)
Land perceived of average quality	0.012	0.035	-0.017
	(0.037)	(0.040)	(0.033)
Adults present	0.010	0.017	0.036**
	(0.016)	(0.016)	(0.015)
Age of household head	0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)
Sex of household head (1=male)	0.032	-0.013	0.017
	(0.041)	(0.047)	(0.041)
Head lower primary (year 1 to 4)	0.015	-0.000	-0.043
	(0.057)	(0.051)	(0.057)
Head higher primary (year 5 to 7)	0.069*	0.066*	0.078**
	(0.039)	(0.034)	(0.033)
Household owns radio	0.067*	0.089**	0.069*
	(0.040)	(0.044)	(0.037)

[a] Pooled techniques. Technique adoption was asked for a series of 10 techniques; standard errors are adjusted for household clustering, probit marginal effects.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table A3
Effect of individual characteristics on banana output – to Table 6.

Dependent: Log(banana harvest value)	All farmers in network included			Highest productive farmer in network excluded		
	Kinship	Neighbours <300m	Informal insurance network	Kinship	Neighbours <300m	Informal insurance network
	(1)	(2)	(3)	(4)	(5)	(6)
Kibanja area (hectare)	0.671***	0.614***	0.368*	0.697***	0.604**	0.501**
	(0.213)	(0.224)	(0.208)	(0.236)	(0.243)	(0.246)
Land perceived of high quality	1.198***	0.895*	1.502***	1.278**	0.934*	1.517***
	(0.439)	(0.507)	(0.403)	(0.484)	(0.533)	(0.437)
Land perceived of avg quality	0.636*	0.720**	1.236***	0.483	0.659*	1.194***
	(0.340)	(0.332)	(0.304)	(0.371)	(0.348)	(0.334)
Adults present	0.327	0.516	0.412	0.190	0.370	0.228
	(0.356)	(0.344)	(0.346)	(0.377)	(0.372)	(0.376)
Age of household head	0.149	0.461	0.002	-0.004	0.358	-0.266
	(0.469)	(0.497)	(0.444)	(0.493)	(0.512)	(0.467)
Sex of household head (1=male)	-0.241	-0.416	-0.336	-0.151	-0.343	-0.298
	(0.282)	(0.284)	(0.257)	(0.307)	(0.302)	(0.281)
Head lower primary (year 1 to 4)	1.026**	1.254***	0.546	1.017**	1.157**	0.532
	(0.447)	(0.423)	(0.432)	(0.475)	(0.438)	(0.447)
Head higher primary (5 to 7)	-0.499	-0.341	-0.432	-0.586	-0.350	-0.561
	(0.385)	(0.382)	(0.344)	(0.413)	(0.394)	(0.363)
Household owns radio	0.537	0.349	0.116	0.504	0.304	-0.089
	(0.347)	(0.320)	(0.284)	(0.380)	(0.331)	(0.311)

* significant at 10%; ** significant at 5%; *** significant at 1%.