CHAPTER 7
MEASUREMENT AND ESTIMATION PROCEDURES

The solution to the optimisation problem in Chapter 5 defines a process of knowledge accumulation, indirect risk and the consequent choice of a management technology. This chapter provides a link between the production behaviour described in Chapter 5 with empirical estimation approaches. The chapter is divided into two parts. The first part discusses the estimation procedure, describing the econometric models used to analyse the use of management practices as well as household participation in associations and private social networks. The second part defines the empirical variables, describes their measurement and hypothesized effects, and puts them into the broader context of the existing literature.

7.1. Econometric estimation procedure

The selection of the empirical estimation approach was dictated by the nature of the data in the sample and the formulation of the economic variables in Chapter 5. The data contained about 20 management techniques farmers could use in the management of their banana plants. Some of these techniques were traditional in the communities and probably easy to apply while others were new and required special knowledge. Still others required additional inputs aside from labour. More importantly, the use of these techniques is a heterogeneous variable that can be measured in multiple ways, including frequency of use, a discrete (zero-one) variable, or a proportion. Constructing one variable or index that combines all of the techniques, or aspects of the techniques, used at the household level is not straightforward.

The econometric approach to estimating the use of banana production management practices consists of a number of steps. First, the 20 techniques were categorized into two major groups according to whether the technique was recommended primarily for soil fertility management or for mat management (also known as sanitation practice). The soil fertility management category contained eight techniques, while the other 12 techniques fell into the mat management category. The eight techniques in the soil fertility management category were aggregated into two soil fertility management technologies (i.e. mulching and manure application technologies) using factor analysis. The estimation results from the factor analysis also provided information that was useful in identifying other necessary statistical tests to be done on the data before final estimation and hypothesis testing.

The 12 techniques in the mat management category were further subdivided into three groups, namely planting technologies, post-planting but pre-harvest technologies and post-harvest technologies. Based on the results of the descriptive statistics, some techniques were discarded from the econometric estimation because of lack of variation, due either to near universal adoption or near universal non-adoption. Techniques excluded from analysis due to near universal adoption include: de-leafing and de-sheathing (both post-planting but pre-harvest management techniques), while those discarded due to near universal non-adoption were hot water treatment and pest and disease resistant banana varieties. Other techniques such as weevil trapping, corm cover and corm removal proved difficult to measure and hence were not included in the analysis. Ultimately, three mat management practices were included in the analysis: corm paring (planting technology), de-suckering (post-planting but pre-harvest technology) and post-harvest pseudo-stem management (post-harvest technology). Pseudo-stems have two components: the lower pseudo-stem and the upper pseudo-stem[1]. In the next subsection a summary of the data reduction method is presented. This is followed by a description of the empirical model and the procedure used in estimation.

7.1.1. Data aggregation, reduction and factor analysis

Data collected from surveys often come in the form of many correlated variables that are difficult to work with individually in statistical analysis. In the present study the survey data contained eight different types of organic materials used in the implementation of soil fertility management, i.e. (1) mulching with grass; (2)

[1] Bananas are normally harvested about one meter above the ground, the part that remains being what is here referred to as the lower pseudo-stem and the part that is cut off during harvesting being referred to as the upper pseudo-stem. After the fruit is cut off, the lower part of the pseudo-stem is also cut off and this is called stumping, while the upper part of the pseudo-stem can either be peeled or chopped up to destroy the breeding grounds of pests and also to facilitate rapid decomposition to recycle nutrients taken up by the plant during its growth. The whole process is what is here referred to as post-harvest pseudo-stem management.

mulching with crop residues; (3) mulching with kitchen residues; (4) addition of cattle manure; (5) goat manure; (6) pig manure; (7) poultry manure; and (8) composting of homestead-type refuse to make manure (hereafter referred to as compost manure).

Factor analysis, with a principal component option was first applied to the eight soil fertility management practices to identify the latent variables that characterize the association between the original variables and to determine whether they could be represented by a small number of components. Based on the criterion of an eigen value greater than unity, the eight soil management practices were grouped into four independent packages according to four unobserved factors (also called latent variables). The results of the factor analysis are summarized in Table 7.

Table 7. Rotated factor loadings of the soil fertility management practices on the four latent variables (indices)*

Variable	Index 1	Index 2	Index 3	Index 4
Crop residues	0.038	0.656	-0.007	0.472
Grass mulch	0.015	0.917	-0.034	-0.226
Kitchen residues	0.841	0.102	0.075	0.029
Goat manure	0.861	0.017	0.193	-0.043
Pig manure	-0.120	-0.013	0.740	0.121
Cattle manure	0.305	-0.054	0.695	-0.096
Poultry manure	0.777	-0.112	-0.307	0.066
Composted manure	0.010	-0.120	0.024	0.939
% Variance explained by factor	28.260	17.870	14.060	12.850

* Interpretation was based on a factor loading of $\geq$ 0.5

The four latent variables explain 73% of the total variance in the adoption components. The latent variables are ordered such that the first latent variable explains the high variation in the data, while latent variables that explain less than one variation are considered as less important and are not included. Next, the latent variables are interpreted depending on the association between them. In the context of the present analysis, four latent variables were generated from the eight soil fertility management practices and were respectively interpreted as (1) traditional technology, (2) mulching, (3) manure application and (4) composting techniques, based on the original variables.

Index 1, interpreted as the traditional technology, explains about 28% of the variance in the eight management practices. Organic household refuse (kitchen residues, goat manure and poultry manure) was highly correlated with this latent factor. All three management practices had positive effects, suggesting that the application of kitchen residues, goat manure and poultry manure are influenced by similar unobserved variables in the same way. These materials are collected in mixtures from the homestead as part of the cleaning activities and spread between mats to control weeds, though some farmers reported that this use was not deliberate. Farmers are advised to compost the household refuse and other organic materials before applying them to banana plantations to facilitate rapid decomposition and avoid the problem of scotching (Tushemereirwe et al., 2003). The technique of composting household refuse before applying it to banana plants was highly correlated (heavily loaded) with index 4. The technique of composting household refuse before applying it to banana plants instead of applying it directly seems to be used independently of other soil fertility management practices.

Index 2 consists of mulching techniques using the organic materials that are gathered from sources other than the banana crop (i.e. crop residues and grass). The use of this type of mulch material in banana production involves the costs of gathering, transportation and application, which reduce the returns from banana production, especially when the transaction costs to access markets are high. This factor explained 18 per cent of the total variance in the use of soil fertility management practices. Both types of mulching materials have coefficients with positive signs. Thus, factors that increase the use of grass mulching are also likely to increase the use of crop residues. Organic fertilizers from animals that are rarely kept in the homestead (cattle and pigs) were heavily correlated with factor index 3. These organic fertilizers also involve costs of access, transportation and application that may limit their use.

The second goal of factor analysis is to summarize the variables contained in the data set in a compact manner so as to be able to relate them statistically with other variables of interest. The application of factor analysis to the correlated variables clusters them into groups according to the latent variables underlying the observed correlation between the variables. The latent variable represents a linear combination of the original variables that captures most of the information in the original variables.

Let $v = \sum_{i=1}^{n} ac_i$ where a_i is a vector of the weights that are mathematically determined to maximize the variation of the linear composite with the original variables and c is a vector of the observable variables in the data set. The sum of the squared weights is constrained to equal one in order to maximize the variation in the composite variables $\sum_{i=1}^{n} a_i^{\ 2} = 1$.

Based on the results of factor analysis, different types of organic materials (c_i) that correlated highly (factor loading ≥ 0.5) with each latent variable were combined by simply adding the banana areas (measured during the mat count) under each organic material. These organic materials were typically applied in separate portions of the banana plantations and hence aggregation did not cause any serious measurement errors. The use of mulching technology is defined as the practice of applying crop residues (non-banana crop residues) or grass to mulch banana plantations. Similarly, the use of manure technology is defined as the practice of applying animal waste (from cattle and pigs) or composted manure to the banana plantations.

7.1.2. Econometric modeling and estimation of banana production management decisions

The decision to use and the extent of use of an improved banana production management technology represent two decisions, although they may be simultaneous in time. The household takes a decision on whether to use the improved management technology or not. Conditional on the decision to use the improved management technology, the household decides on the extent of use of the technology. The econometric approach used to estimate decisions regarding the use of banana management practices can be linked to the theoretical model through an index function model involving decisions about whether or not to use the technology and how much of it to use (Greene, 2000). Denote y^* as a vector of the unobserved demand for the improved management technology, as follows:

$$y^* = \alpha' Z + v \qquad v \mid Z \sim N(0,1) \qquad (31)$$

Z represents a vector of explanatory variables cast on the right-hand side of equation (30); v is a vector of unobserved heterogeneity; and α is a vector of the parameters to be estimated. At the time of the survey demand had not been observed for some households and hence the structural equation cannot be estimated. Instead, a reduced form is estimated and the focus is on two management decisions, the discrete decision (to use or not to use) and the extent of use.

The household's decision to use the improved management technology is only observed when the latent variable exceeds a threshold value. Suppose the choice to use an improved management technology is observed when the latent variable is greater than zero and remains unobserved when the value of the latent variable is unknown. The reduced-form equation of the choice of the improved management technology can then be specified as:

$$\begin{aligned} y &= 1 \quad && if \quad y^* > 0 \\ y &= 0 \quad && otherwise \end{aligned} \qquad (32)$$

We cannot observe the land allocation to the improved management technology for cases where $y = 0$, but only for a subset of the population for which $y = 1$. Data on the extent of use (δ^*) of the improved management technology is missing for those households who did not choose to use the improved management technology. Missing data for a set of explanatory variables leads to a censoring of the demand for the improved management practices for which $y = 0$ (Maddala, 1983; Greene, 2000; Wooldridge, 2002).

A Tobit regression model has been widely used in estimation when the dependent variable is observed within a limited range (Greene, 2000). Underlying the Tobit model is the assumption that the coefficients on the probability and extent of adoption are the same (Greene, 2000). Thus, the Tobit model fails to separate the two decisions that characterize the adoption of a divisible technology. The decision to use and the extent of use are also likely to be influenced by different factors (Wooldridge, 2002).

To test whether the Tobit model is a suitable representation of the processes affecting the use of improved banana management technology, Probit, Tobit and truncated regressions were estimated for each of the three technologies. The null hypothesis of equal coefficients was tested using the likelihood ratio statistic, where the restricted regression is the Tobit model and the unrestricted regression is the combined Probit and truncated regression. The results are summarized in Table 8. For most of the management practices the statistical significance of the test statistic leads to rejection of the null hypothesis that the coefficients are equal. The data therefore support separate estimation of the probability of use and extent of use decisions.

Table 8. A likelihood ratio test for the null hypothesis that the coefficients on the two management decisions are the same

	Value of the log likelihood function			Likelihood ratio
Management practice	Tobit model	Probit model	Truncated regression	test (P-value)
Mulching	-90.358	-147.680	134.690	0.000
Manure	-110.900	-163.500	91.900	0.000
Post-harvest pseudo-stem management	-197.620	-98.070	-83.820	0.219

The reduced-form model describing the banana area share allocated (the extent of use) to improved management technology in the population is specified as:

$$\delta^* \mid y=1 \quad = \beta' X + \varepsilon \tag{33}$$

$$y=1 \quad if \quad y^* > 0$$
$$y=0 \quad otherwise$$

As defined above, δ^* is the optimal observed area share of the bananas under the technology, X is a vector of explanatory variables, and $E(\varepsilon \mid X)=0$; it is assumed that the unobserved heterogeneity in the vector ε is uncorrelated with the exogenous variables.

The two management decisions can be estimated in two stages. The first stage uses a standard Probit, as specified in equation 31, and estimates the probability of using a management practice on the whole sample. In the second stage, an OLS regression can be used to estimate the extent of management on a sub-sample with non-zero technology use.

The fact that the demand for the improved management technology is now observed for a sub-sample of the population can create a sample selection problem that can result in inconsistent estimates in the extent of management equations (Maddala, 1983; Wooldridge, 2002). A Heckman two-step estimation procedure was used to test and correct for sample selection bias in the data. Although the share ranges from 0 to 1, the sample data indicate that all households have an extent of use of less than one. Thus, a model that accounts only for censoring at zero was applied.

The results of the first stage, described in the previous section, are identical to the Heckman first step. In other words, in the first step of the Heckman procedure, the choice of whether or not to use an improved management practice, is estimated for the full sample using a Probit model as specified in equation (31). Regression of the binary response variable (y) on the explanatory variables gives predicted estimates used in the computation of an inverse Mills ratio:

$$\lambda(\alpha^{'} Z) = \phi(\hat{\alpha}^{'} Z) \Big/ \Phi(\hat{\alpha}^{'} Z)$$

The inverse Mills ratio is a non-linear function of the density and distribution defined over the Probit estimates. The statistic captures the information related to the sample selection. The computed inverse Mills ratio is then included in the extent of management practices estimation in the second step. The statistical significance of the coefficient on the $\lambda(\hat{\alpha}^{'} Z)$ implies that the two error terms are correlated and hence confirms the presence of selection bias (Wooldridge, 2002). When the null hypothesis is not rejected, then the inverse Mills ratio should be included in the second stage estimation of the extent of use of management practices to correct for standard errors. When the null hypothesis is rejected, it means that a sub-sample of non-zero use of the improved management technology is representative of the population and the extent of use of the management practice can be estimated using OLS.

Hypothesis tests do not support the presence of sample selection bias in estimating the extent of use of mulching and post-harvest pseudo-stem management practices, but support it in the case of the use of manure. The test results imply that in the mulching and post-harvest pseudo-stem management equations, the sub-sample of households with non-zero use is representative of the population (Wooldridge, 2002). Consequently, the extent of use of these management practices was estimated by OLS regression on a sub-sample with non-zero use, while a Heckman model was used to estimate the extent of manure application.

According to the economic analysis developed in Chapter 5, the decisions regarding the use of banana management practices are also conditioned on farmers' perceptions of the biotic and abiotic problems, a random variable. Farmers' perceptions about these factors are influenced by the same variables that influence the use of improved management technology. Thus, we are dealing with a simultaneous equations model that is a function of exogenous variables, predetermined variables and an error term. Decisions about the use of banana management practices can be estimated as a function of direct measures of perception or by substituting for perception, using exogenous factors in the perception equation. The use of the direct measure of perception in the management use equations creates the problem of endogeneity. The observed indicator of perception is correlated with the error term ε of the demand equation, thus rendering the OLS estimates inconsistent (Greene, 2000; Woodridge, 2002). Consistent estimates can be obtained by using a two-stage least-squares estimator (2SLS) to correct for endogeneity (Wooldridge, 2002). If correctly specified, this estimation procedure will yield estimators with greater asymptotic efficiency than are attainable by the limited information method (Greene, 2000). However, this approach requires extensive data, which in most cases are not available. In addition, the full information method is complex when the null hypothesis of sample selection bias has not been rejected. Since the main focus of this study was to test the effect of social capital while controlling for other factors, the demand for improved banana management technology was estimated as a function of all exogenous variables cast on the right-hand side of equation 30, expressed as a reduced form.

The final statistical consideration is the possible simultaneity in adoption decisions for these management practices. The results of the factor analysis did not support the simultaneity of the manure and mulching adoption decisions. Even when the use of the two types of practices is determined independently, it is possible that the unobserved heterogeneity in the technology demand equations is correlated. In this case**,** statistical efficiency can be improved by joint estimation. Nonetheless, estimating the manure and mulching decisions jointly was not considered to be worthwhile since the set of explanatory variables is identical in all equations (Greene, 2000), implying no gains in statistical efficiency. Therefore each technology was treated separately.

7.1.3. Econometric estimation of household participation in associations and private social networks

In the theoretical analysis developed in Chapter 5, social capital was treated as a fixed variable. It was assumed that at the time of making the choice of a crop management technology, a certain amount of social capital stock that interacted with decision-making had been accumulated over the previous period. In this section, that assumption is maintained and the observed memberships in associations and the stock of private social networks resulting from household investment decisions is analysed. The purpose of this is to discern what characteristics of the households and/or the community explain the differences in these two forms of social capital.

Social capital accumulation is conceptualised borrowing ideas from the model of Glaeser et al. (2001). Participating in an association or joining a private social network is a decision-making process that involves comparing costs with benefits. Participation can be conceptualised as a cost-benefit decision that takes into account present expenditure in terms of time and other goods, as evaluated against immediate and future expected returns. It is reasonable to expect that when an individual contemplates joining a social network and investing in social interactions, he/she understands that there are benefits as well as costs involved. Benefits can take the form of informal credit, information, other material goods and/or emotional support. An example of a less tangible benefit is the direct happiness or social approval from others that an individual obtains from participating in a community association such as a burial society or religious organization. Costs may be in terms of utility loss from consumption foregone because time and valuable resources spent on socializing reduce the resources available for work. The individual views these prospects within a cost-benefit framework.

For different associations there may be different benefits and costs, which could motivate an individual to participate in various associations. Hence, analysing participation in associations involves the estimation of a binary decision whether or not to join a particular association and a decision as to the number of associations to join, a discrete-count variable. The number of associations a household participates in reflects the intensity of the household's social capital. Both decisions are based on the expected net benefit, which is unobservable to the researcher but assumed to be positive when the household decides to participate in the association. The econometric analysis of the two decisions is described below.

7.1.3.1 Decision to join an association

Denote Y_{ik}^{*} as the expected unobserved net benefit of participation in a local association. The decision as to whether to participate in any association k or not, is defined as a binary outcome (P) of an observed latent variable (Y_k^{*}). The latent variable underlying this decision is a linear function of its observed (W_k) and unobserved (ε_k) determinants, such that:

$$\begin{aligned} P_{ik} &= 1 \quad if\ Y_{ik}^{*} = \gamma^{A} W_k + \varepsilon_k \quad > 0 \\ P_{ik} &= 0 \quad otherwise \end{aligned} \tag{34}$$

Probit[2] estimation is appropriate for estimating models with such binary dependent variables. The model for the membership in each association was estimated separately.

[2] Logit estimation is also generally appropriate for analyzing binary-response data, and under standard assumptions about the error term there is a priori reason to prefer probit estimation to logit estimation (Greene, 2000).

7.1.3.2 Intensity of group membership and private social networks

The number of memberships held by a household or the number of trusted people to whom a household is connected can be modelled as a series of discrete household decisions that sum across an aggregation of choices to a Poisson distribution. The Poisson model is a non-linear specification and estimates the effect of independent variables on a scalar dependent variable. The density function for the Poisson regression is:

$$f(S_i \mid W) = \frac{\exp(-\mu)\mu^{S_i}}{S_i} \tag{35}$$

$$(S \mid W) \sim Poisson(\mu) \quad or\ NegBin \quad (\mu, q)$$

where the mean parameter (μ) is a function of explanatory variables that influence the household decision to participate in associations or private social networks, expressed as vector W and a parameter vector, γ^D. Included in the vector, W, are household and community-specific factors. The descriptive statistics and measurement of these variables are discussed in section 7.2.5. The scalar, S_i, is the dependent variable representing the household membership in associations or household density of private social networks. For a Poisson distribution, $E[S_i \mid W] = \mu = \exp(\gamma^D W)\ and \quad S = 0,1,2....s$, taking an exponential of ($\gamma^D W$) causes the expected count μ to be positive, which is required in a Poisson distribution (Long, 1997).

The validity of the Poisson model hinges on the assumption that the conditional mean is equal to the variance. In other words, a Poisson distribution is a single-parameter distribution with a mean equal to the variance $E(S \mid W) = Var(S \mid W) = \mu$. In most common applications, however, the conditional variance is greater than the conditional mean. A Negative Binomial regression that accounts for unobserved heterogeneity was fitted in order to test for over-dispersion. A Negative Binomial Regression model was also used to relax this assumption, allowing for over-dispersion in the data, such that $E(S \mid W) \neq Var(S \mid W) = \mu + \sigma = q$ (Greene, 2000). The statistical significance of the over-dispersion parameter against the null hypothesis of equi-distribution is rejected in the case of estimating the density of the private social network, implying that data on the number of trusted friends may not exhibit a Poisson distribution. A Negative Binomial regression was used to estimate the density of private social networks, while a Poisson model was used to estimate the density of membership in associations.

7.2. Definition and measurement of variables

In this section the empirical definition and measurement of the variables used in the empirical estimation are described. The section first discusses the methodological approaches to adoption studies. This is followed by the definition of the dependent variables used in the estimation of the use of management practices and social capital, elaborating on how each was measured. Then the independent variables are defined and their hypothesized effects and measurement are described.

7.2.1. Econometric approaches in technology adoption studies

Technology adoption studies can be categorized into three groups according to the type of data used: time series studies, cross-sectional studies and studies that use panel data (Besley and Case, 1993). Time series studies focus on the aggregate measures and the rate at which the technology was diffused in a specific community or region. Typical examples of such studies are Griliches (1957) and Mansfield (1961). This approach provides an insight into the community or regional characteristics and the technological attributes that influence the rate of technology adoption (Rogers, 1995), but is limited as to what it can say about the dynamic processes that influence technology adoption and diffusion.

Temporal studies that use panel data sets in which the same decision-making unit is observed over a period of time have been used to overcome such limitations. These studies capture both the dynamic processes that influence technology adoption and the impact of the technology on the adopting households and income distribution (examples are Foster and Rosenzweig, 1995), but they are rarely used because of the high cost of data collection. Instead, most studies on technology adoption make cross-sectional observations to classify the population into adopters and non-adopters, although such studies have been the subject of criticism in recent years (Besley and Case, 1993), thus drawing attention to methodological issues regarding cross-sectional studies of technology adoption.

Besley and Case (1993) classify the cross-sectional studies on technology adoption into two groups: (1) studies that take a snapshot to analyse the impact of farm and farmer characteristics on technology adoption; and (2) studies that use recall methods to go back over the history of technology adoption decisions. The second approach has advantages over the first in that it allows the dynamic process to be modelled in a manner that gives an insight into the importance of the impact of the previous state of nature on the current adoption decisions. However, as Besley and Case (1993) noted, this approach, while an improvement, still has limitations because adoption decisions may influence some explanatory variables, thus rendering them endogenous.

These limitations can be overcome by using recall to obtain information on the explanatory variables that are expected to influence adoption decisions before the first adoption decisions were made, but if the period of adoption goes far back in history, then the reliability of recall for such data is questionable. Another approach sometimes used to overcome endogeneity in the explanatory variables suspected to be influenced by previous adoption decisions is to estimate a system of simultaneous equations. Recent examples of such methodological approaches were applied in Negatu and Parikh (1999) while analysing the impact of perceptions on adoption decisions. Other researchers that have focused on new methodologies to deal with issues of the endogeneity and simultaneity of adoption decisions and incomplete technology diffusion include, amongst others, Smale and Heisey (1993) and Dimara and Skuras (2003). Smale and Heisey (1993) modelled adoption as three simultaneous choices: the choice of whether to adopt the component of the recommended package, the decision as to how to allocate different technologies across the land area, and the decision as to how much of certain inputs, such as fertilizer, to use.

Like most other studies, the present study uses cross-sectional data. As the descriptive information suggests, most of the management practices investigated had been in the communities long enough for characterizing adopters and non-adopters to yield reliable information regarding the adoption of these management practices (Appendix A).

Programme selection bias is also a common problem, related to inadequate definitions of the "counterfactual". This problem is typically addressed by the application of a treatment model, which implies a sampling strategy that includes a control and a treatment group. In the present study, the sample selection problem has been accounted for in the sampling frame for the study by allocating the sample to "exposed" and "non-exposed" areas.

7.2.2. Dependent variables for adoption models

Defining technology adoption at the individual level is a complex matter that depends on the nature of the technology, the local context and the research questions being answered (Dossy, 2003). Technology can be defined in terms of a discrete structure (0,1) when the technology in question is used exclusively, as in case of non-divisible technologies. When technologies can be adopted partially, a continuous variable is a more appropriate measure of adoption (Feder et al., 1985). Empirically, continuous decisions have been measured in terms of proportion, scale or intensity of use. Sometimes more than one continuous measure is used to reveal important information about the adoption behaviour (Smale and Heisey, 1993; Gebremedhin and Swinton, 2002). For example, Gebremedhin and Swinton (2002) found that the factors affecting the proportion and intensity of soil conservation were different.

Measuring the use of improved banana management technologies is even more complex due to the fact that the actual technology is made on the farm and there is no standard measure of adaptations. This perhaps means that modifications also vary across farms, which makes it difficult to establish a measure of the variants of the technology that can be generalized across the sample. Further complexity regarding the measurement of the intensity of use per mat is associated with the variability of mat sizes within the banana plot. Considering this complexity, the proportion (in terms of mat share) of use is a simpler measure for representing the extent of management with a specific technology. Hence the demand for an improved banana management practice is defined as the proportion of mats managed using a particular practice.

The term "use" rather than "adoption" is used in recognition of the complexity of defining adoption either in terms of a decision (discrete [0,1], proportional use [mat share], scale of use [number of mats], level of choices [per mat] and time of use [testing or farmer experimentation as compared to long-term use]). Hence the term "use" as used in the present study represents behaviour that could constitute either experimentation with the technology or final adoption after confirmation of the utility of the technology.

Although most banana management practices are continuous by nature, corm paring and de-suckering proved difficult to measure quantitatively. As such, for the purposes of this study, these practices are defined in discrete terms. Definitions of these and other management practices included in the analysis are presented below. A summary of the descriptive statistics on the respective adoption-dependent variables is presented in Table 9.

7.2.2.1 Corm paring

The use of "corm paring" technology is defined according to its discrete structure, i.e. 0,1 is measured as one if the farmer reported that the technology is used and zero if not.

7.2.2.2 De-suckering

De-suckering was measured as the average number of plants per mat. Farmers with an average of four or fewer plants per mat were considered to be adopters of de-suckering, while those with an average of above four plants per mat were categorized as non-adopters[3]. The variable is coded as binary. De-suckering could have been measured as an integer to allow more variability in the sample but this was made difficult by the variability of the mat plant population within a plot and averaging over the plot changed the variable from an integer to a continuous variable. Since the focus of the study is on examining deviations above the recommended number, irrespective of how much the farmer deviated from the recommendation, a binary indicator is appropriate.

7.2.2.3 Mulching, manure application and post-harvest pseudo-stem management technologies

The use of mulching, manure application and post-harvest pseudo-stem management were observed as continuous variables. Respectively, the use of mulching or manure application is defined as the proportion of mats managed with organic mulch or manure. Similarly, post-harvest pseudo-stem management is a continuous variable defined as the proportion of the managed pseudo-stems from which the fruit has been harvested and the technique involves either stumping, splitting or chopping. "Total pseudo-stems" is a count of all lower pseudo-stems and upper pseudo-stems left after harvesting for a period of one year. The use of these three management practices was measured for a period of 12 months (equivalent to one crop production cycle).

Table 9. Summary statistics of the adoption-dependent variables

Variable	Definition	Mean	SD
Corm paring	A binary indicator = 1 if the household reported use of corm paring before planting and = 0 if corm paring was not used	0.233	0.423
De-suckering	A binary variable = 1 if the average number of plants per mat is ≤ 4 and = 0 if otherwise	0.450	0.498
Use of mulch (δ_1)	Share of banana mats grown under mulching	0.199	0.260
Use of manure (δ_2)	Share of banana mats grown under manure	0.119	0.233
Post-harvest pseudo-stem management	Proportion of post-harvest pseudo-stems harvested that were managed either by stumping, splitting or chopping	0.320	0.286

[3] The recommended banana management technique is to allow three plants per mat but two or four may be left on a mat to adapt to the soil conditions.

7.2.3. Dependent variables for social capital models

Social capital is an unobservable asset, which is empirically measured through the use of proxies. In the present study, associations and private social networks were used as proxies for social capital. This subsection presents the definition of and measurement of the variables used to assess participation in associations and private social networks. The descriptive statistics for each variable are presented in Table 10.

7.2.3.1 Participation in associations

The decision to participate in associations is defined as binary (equal to one, if the household has membership in the association in question, and zero if not).

7.2.3.2 Intensity of participation in associations

The intensity of participation in associations at the household level is defined as the total number of memberships in associations held by household members. This is computed as the sum of the number of memberships held by household members.

7.2.3.3 Intensity of private social networks

The intensity of private social networks is defined as the number of trusted friends the household can rely on for help in case of any problem (i.e. the number of friends household members can talk to intimately, approach about any problem or with whom they can freely share a family secret). Conceptually, this measure is related to the proxies used by Godquin and Quisumbing (2005). The difference is that these authors disaggregated the measure by the hypothetical problem. Here, it was not considered important to disaggregate the measure by the hypothetical problem, since individuals may not have separate networks for separate problems or types of trauma.

Furthermore, the definition of private social networks used in the present study also differs from that of other definitions of social networks in that in this case relatives are excluded from private social networks. Relatives are considered to constitute a “given” social capital whose formation may be beyond the influence of the decision maker. It therefore constitutes an initial stock of social capital that could be included in the estimation as an explanatory variable rather than as an endogenous dependent variable.

Table 10. Descriptive statistics of the social capital dependent variables

Variable	Definition	Mean	SD
Membership in any association	A dummy variable = 1 if the household participates in at least one association and 0 if not	0.747	0.435
Membership in social association	A dummy variable = 1 if the household participates in either religious, culture-based or burial associations and 0 if not	0.610	0.488
Membership in informal saving and credit association	A dummy variable = 1 if the household participates in any informal credit and saving associations and 0 if not	0.222	0.416
Membership in agriculture-based association	A dummy variable = 1 if the household participates in any agriculture-based association and 0 if not	0.175	0.380
Household intensity of association	Total number of memberships in associations held by household members	2.232	2.256
Number of friends	Number of people the household members talk to intimately, with whom they share family secrets or who can be approached for help in case of any problem and to whom they are not related by blood	14.958	13.719

The household private social networks were elicited using a position-generator technique discussed in Lin (1999). The technique starts with representative positions in the society and the individual is asked whether he/she has ties to people in each position. By eliciting ties to people in each different position individually this method avoids the biases inherent in ego-centred network mapping methods. Ego-centred network mapping elicits a list of ties from the individual together with the relationships between the individual and the tie (Lin, 1999). Ties elicited using the ego-centred network mapping method are biased towards the stronger ties and the network may be under-reported (Lin, 1999). Both ego-centred network mapping and the position-generator technique emphasize the measurement of social capital as a resource embedded in social networks.

An alternative strategy for measuring social network capital focuses on the location within a social network as an indicator of the individual's social capital. Empirical measurement of network locations is accomplished by complete mapping of the network (Broeck, 2004). The advantage of this method is that it allows a detailed and complete analysis to be carried out of all network locations and embedded resources. However, the method assumes that a network has a defined boundary. This is useful when the focus of the study is on examining networking within a small location (such as a village) or organization, but not in the case of extensive surveys.

7.2.4. Independent variables used in adoption equations

Conceptual variables included in the analysis of the use of banana management are represented by the reduced-form equation and their operational definition was guided by the existing literature. Banana production is a semi-subsistence practice across much of the survey domain, with uneven access to markets and market participation, consistent with the non-separable case of the household model. When the consumption and production decisions are non-separable, the effects of comparative statistics are ambiguous. Thus, hypothesized effects are motivated by related adoption literature and previous information concerning banana production in Uganda.

The reduced-form equation for the improved management technology cast in equation 30 in Chapter 5 indicates that the banana area share allocated to an improved management practice is a function of the village wage rate for unskilled labor (w), the banana market price (P^B), household characteristics (Ω_{HH}), market characteristics (Ω_M), farm characteristics (Ω_F), exogenous income from private assets (I), formal information diffusion (Ω_D), farmer experience with the technology (τ), and different forms of social capital (Ω_{SS}). In this section, the empirical definitions and measurement of these variables are discussed. The descriptive statistics on each explanatory variable and the hypothesized effects thereof are summarized in tables presented in the text.

7.2.4.1 Household characteristics (Ω_{HH})

A number of household characteristics are hypothesized to directly or indirectly influence the choice and extent of use of the improved banana management technology. Some of these characteristics are specific to the farmer, who is defined in the present study as the primary production decision maker. They include age, gender and human capital. Human capital is represented by formal education, measured as the number of completed years of schooling. Education was identified in earlier adoption studies as an important household characteristic in adoption decision-making processes (Feder et al., 1985; Feder and Umali, 1993). Higher education is associated with the capacity to understand technical aspects related to the new technology but may also increase the opportunity cost of labour, which could reduce the use of labour-intensive management practices. Older decision makers are expected to discount the future heavily, implying that age is associated with low investment in techniques whose benefits and costs are far separated in time (Shiferaw and Holden, 1998). Gender is also included to assess whether there are any gender disparities in the use of banana management practices. Some practices such as de-suckering require greater physical strength for their implementation that may not be possessed by females, which could result in gender disparities in the use of management practices.

Other household characteristics included in the analysis were: wealth, household size and the dependency ratio. The effect of these factors on management decisions depends on the nature of rural market imperfections (Pender and Kerr, 1996). When labour markets are imperfect, households endowed with family labour may be more able to meet the high labour demand of improved banana management technology (i.e. mulching, pseudo-stem management and manure application) than their counterparts with smaller family labour endowments. Similarly, given the missing markets for organic fertilizers, households endowed with the assets (such as land and livestock capital) that produce these materials will be able to invest more in managing the soil fertility of their banana plantations.

Landholding (also referred to as farm size) was measured as the total number of per capita cultivable hectares possessed by the household. Data show that landholding size is positively correlated with the cultivated area. This implies that the supply of crop residues from other cultivation practices or grass needed to implement the mulching technology will be positively correlated with the landholding size. Since these fertilizers are produced on the farm as by-products of other farm activities, a larger landholding size should increase the capacity of the household to implement mulching technologies. Hence, landholding size may be expected to correlate positively with the demand for mulching technologies. Even when households have access to other land through hiring or borrowing, the high cost of transporting the residues to their family plots may impede their use in banana production.

Landholding can also act, through its influence on perceptions, to influence banana management decisions in the opposite direction. Boserup (1965) hypothesizes that increasing population pressure stimulates the use of land intensification techniques. This means that while more extensive landholdings per capita enable households to engage in crop production, which may result in organic materials for mulching, they are nevertheless associated with less pressure on the cultivable land. Low pressure on land, on the other hand, may suppress the perception of soil fertility problems, thus reducing the demand for soil fertility management practices.

The effects of owning livestock also appear to be ambiguous. The possession of livestock reduces the cost of access to animal manure, which may result in greater use being made thereof. On the other hand, the accumulation of livestock may imply a shift away from crop production and consequently a reduced pressure on the land, which could lower the perception of soil fertility problems and hence the use of practices related to soil fertility management. The number of livestock units is converted following Nkonya et al. (1998), as demonstrated in Table 11.

Table 11. Descriptive statistics of household characteristics (Ω_{HH})

Variable name	Description	Expected effect	Mean	SD
Income (I)	Net transfers from household private assets	+	106 087.00	1 667 922.00
Age	Age in years	-	43.08	15.67
Gender	Gender (1 = male, 0 = female)	+/-	0.55	0.51
Education	Completed years of schooling	+	4.56	3.86
Household size	Total household members	+	5.87	2.77
Landholding	Total hectares of cultivable land per household member	+/-	0.37	0.68
Livestock unit	Sum of cattle units (0.8), sheep (0.4), pigs (0.4) and goats (0.4), divided by age household head	+	0.03	0.04

Distortions in output markets encourage self-sufficiency in that output, implying that an increase in the consumer-worker ratio (dependency ratio), which increases the household consumption demand, will stimulate investment in production. However, when insurance and labour markets also fail, as is the case in rural Uganda, a higher consumer/worker ratio may increase the risk of starvation, which could limit investment in labour-intensive activities. Hence the effect of the dependency ratio on the use of improved banana management practices cannot be determined a priori.

The expenditure constraint in the estimation is represented by the total exogenous income received by the household as a net transfer in the form of interest from private assets (I) and income in the form of gifts and remittances from social capital (Ω_{SS}). Detailed information on the definition and measurement of bilateral transfers that accrue to the household from its social network are presented in section 7.2.4.5.

The exogenous income from household private assets (I) was defined as cash inflow in the form of rent from buildings or interest on previous investments. Households with access to this type of exogenous income are able to overcome liquidity constraints and can purchase farm implements that will enable them to save on labour or use the cash to hire labour. Hence they will be able to use and apply banana management practices more extensively.

7.2.4.2 Farm characteristics (Ω_F)

The household's decision to use an improved banana management practice and consequent demand for the technology is also hypothesized to depend on farm characteristics such as location, physical land factors and scale of production. Elevation, a sample stratification parameter, is expected to condition both the use of practices and perceptions. At high elevations the soil erosion potential is higher, which could affect farmers' perception of the soil fertility problem and consequent investment in mulching and/or manure, to conserve soil (Ervin and Ervin, 1982). Prior biophysical information also indicates that banana productivity potential is higher in high-altitude areas, which provides further incentives to use improved management practices. The high pest and disease pressure in low-elevation areas implies that the risk from these biotic factors is greater in these areas than in higher areas. The other factor positively correlated with the risk of biotic/abiotic factors is the age of the plantation. Older plantations may be associated with a higher risk of pest/diseases or soil fertility problems because of the length of time these constraints have had to accumulate. However, the age of the plantation could also be related to lower risk due to continued management efforts to mitigate the negative consequences of these constraints. Hence the effect of the age of the plantation is ambiguous.

Physical land characteristics are hypothesized to directly affect perceptions of biotic and/or abiotic factors (i.e. pests, diseases and soil deterioration) but to affect the use of management practices only through their impact on perceptions. These variables include: the slope of the farm, the moisture retention of the soil in the banana plot and the drainage conditions of the banana plot. The physical land characteristics were measured in qualitative form using a subjective measure reported by the farmer (Table 12). The slope of the farm represents the erosion potential, while the capacity of the soil to retain moisture indicates the capacity of the soil to support the high demand of the banana crop for water. Therefore, a low soil moisture retention capacity in banana plots should increase the demand for mulch and manure through its influence on perceptions. On the other hand, the counteracting effect of poor soil moisture retention capacity on the productivity of mulching or manure could lower the incentive to incur costly investments.

Table 12. Descriptive statistics of farm characteristics (Ω_F)

Variable	Description	Expected effect	Mean	SD
Elevation	Elevation at which the farm is located (1 = high 0 = low)	+/-	0.28	0.45
Physical land characteristics				
Slope of the farm	A dummy = 1 if the slope is rated steep and = 0 if otherwise	+	0.78	0.42
Soil moisture retention capacity	A dummy = 1 if the soil moisture retention capacity of the banana plot is rated low and = 0 if otherwise	+	0.21	0.41
Drainage conditions on the plot	A dummy = 1 if the drainage conditions on the banana plot are rated poor and = 0 if otherwise	-	0.39	0.49
Mats	Total number of banana mats grown	+	283.43	334.68
Age of banana plantation	Number of years the banana plantation has been in existence since its establishment	+/-	16.49	20.25

The drainage capacity of the banana plot is another land quality characteristic that may be important in banana production (Tushemereirwe et al., 2003). Bananas grown on poorly drained soils are more vulnerable to leaf spot diseases, which may stimulate perceptions of disease problems. Because of the increased perception of diseases, the decision maker may require more of those management practices that are capable of mitigating the effects of the diseases. However, since there is no effective control measure for these diseases, the likelihood of the diseases occurring may pose an exogenous risk to banana production and could discourage investment in costly banana management practices.

The scale of production (measured by the banana mat count) reduces the fixed cost of information acquisition per unit area, thereby increasing the benefits from adoption (Feder and O'Mara, 1981). Like the landholding size, the scale of production could also act through perceptions to reduce the demand for management practices. Both factors reduce the economic impact of biotic and abiotic factors on the household. However, the two factors work through different mechanisms to increase the demand for management practices, thus justifying the inclusion of both measures of farm characteristics in the estimation.

7.2.4.3 Market prices and related characteristics (Ω_M)

Market access was measured as a village-level variable. The level of infrastructure development, measured as the distance in kilometres from the centre of the village to the nearest paved road, was used as the proxy for physical access to the markets. The direct link between infrastructure development and the use of banana management practices is not clear. Improvements in road infrastructure reduce the costs of physical access to markets for bananas, but also enhance market opportunities for non-agricultural enterprises, thus increasing the opportunity costs of investing in agriculture and labour for banana production. In the latter case, the indirect effects on the use of land-intensive crop management technologies could be negative when road infrastructure improves.

Economic theory predicts that, all other things being equal, a higher market price for bananas will increase the net returns from the higher yields associated with better crop management technology, while higher input prices (e.g. wage rates) would reduce the returns and hence the incentive to use improved banana management technology. Because the banana farm-gate price was positively correlated with the wage rate, a ratio (P^B/w) of the average farm-gate price of bananas to the village wage rate was constructed for the estimation (Table 13). Note that prices were measured as village means to reflect high intra-village correlation relative to inter-village correlation.

Table 13. Descriptive statistics of market characteristics

Variable	Description	Expected effect	Mean	SD
Mean wage rate	Village unskilled mean wage rate (in Uganda Shillings)	-	1 630.00	696.85
Banana farm-gate price	Village mean farm-gate price for bananas (in Uganda Shillings)	+	125.99	50.61
Price/wage ratio (p^B / w)	Average banana farm-gate selling price divided by the average wage rate for hired labour	+	0.10	0.03
Distance	Distance from paved roads (in kilometres) to the centre of the village	+/-	10.69	7.05

7.2.4.4 Information diffusion parameters

The stock of knowledge was conceptualised in Chapter 5 as a shifter of the banana production function, which increases the net returns from banana production. The farmer's level of knowledge also affects the process of forming perceptions. The process of accumulating knowledge is unobservable but knowledge accumulates as a function of experience (τ), formal information diffusion parameters (Ω_D) and social capital (Ω_{SS}) for a given level of education. Information dissemination and diffusion parameters are hypothesized to affect the use of practices both directly and indirectly through perception formation. Formal diffusion parameters included in the estimation are extension and exposure to the new banana varieties. Extension is measured as the number of cumulative contacts with extension educators in the period before the commencement of the study. Exposure to the new banana varieties is a dummy variable =1 if the village was exposed and = 0 if it was not exposed (Table 14). Large, discrete differences in knowledge are hypothesized between villages that have been exposed to the new banana varieties and related information and those that have not. The positive role of extension educators in the adoption of agricultural technologies is well established in the literature (Feder et al, 1985; Feder and Umali, 1993).

Table 14. Descriptive statistics of the parameters of information diffusion

Variable	Description	Expected effect	Mean	SD
Experience regarding mulch (τ_1)	Years of mulching divided by age	+	0.25	0.24
Experience regarding manure (τ_2)	Years of manure use divided by age	+	0.11	0.17
Experience regarding mat management (τ_3)	Years of use of post-harvest pseudo-stem management, de-suckering bananas	+	0.45	0.51
Formal information diffusion (Ω_D)				
Extension	Number of cumulative contacts with extension educators	+	1.15	2.05
Exposure	Dummy = 1 if a village was exposed to the new improved banana varieties and = 0 if otherwise	+	0.50	0.50

Social capital and farmer experience represented informal means of generating or diffusing information. Experience with the technology affects both perceptions of biotic and abiotic factors and the use of management practices. The farmer's experience was measured as the number of years the technology had been used on the farm (corrected for age) (Table 14). The definition and measurement of social capital variables included in the estimation are discussed below.

7.2.4.5. Social capital variables (Ω_{SS})

The conceptual framework developed in Chapter 5 highlights two mechanisms, viz. exogenous income and social learning effects, through which social capital can influence household decisions about improved banana management technology. Because of the importance of social capital in the present study, a number of social capital indicators were included. These are: household density of membership in associations, norms of decision-making, leader heterogeneity in associations and bilateral transfers in the form of labour, cash and durable consumer goods. The descriptive statistics of these variables are summarized in Table 15.

Household membership density is defined as the number of household members who belong to at least one association. Household membership density reflects the household's capacity to acquire information from the social network and the extent to which household decisions are influenced by the decisions of other households. This variable is also expected to influence the use of banana production technology directly since it measures participation in associations engaged in economic activities, which may reduce expenditure constraints on labour use or the acquisition of organic fertilizers (Narayan, 1997). However, the number of household members who join an association can also influence the opportunity cost of time used for banana production. Hence the nature of the effect cannot be determined a priori.

Table15. Descriptive statistics of social capital variables (Ω_{SS})

Variable	Variable description	Expected effect	Mean	SD
Household membership	Number of household members who belong to at least one association	+/-	1.25	1.0
Leadership heterogeneity	A continuous index measuring the degree of leader heterogeneity in terms of livelihood or level of education higher than most group members	+	4.44	0.6
Norm of decision making	A continuous index measuring the degree to which decision making in associations is participatory	+/-	6.11	0.5
Bilateral transfers				
Net labour Transfers	Value (in Ugsh'000) of labour obtained from the social network less the value of labour supplied to the social network	+/-	339.26	8 153.5
Net cash transfers	Amount of cash (in Ugsh'000) obtained from the social network less the amount of cash supplied to the social network	+/-	-1 554.60	163 111.0
Net transfer of consumer durables	Value (in Ugsh'000) of other household items obtained from the social network less the value of those items supplied to the social network	+/-	295.76	39 791.6

Decision-making norms and the heterogeneity of leaders (in terms of livelihood and education) in associations were used as proxies for the characteristics of associations. Decision-making norms measure the group's ability to cooperate, share information while the leaders with higher education and livelihood status depict the opinion leadership in the association and its ability to network with other people beyond the village community. Both decision-making norms and leader heterogeneity are variables that are constructed at the village level.

The concept of "decision-making norms" is defined as the degree to which the members in an association participate in important issues of the association, computed as an additive index from responses to two questions regarding the selection of leaders and the decision making of their associations. Respondents were asked how important decisions were made. Responses ranged from 1 = only leaders participate, to 2 = few members participate, to 3 = all members participate. They were also asked how the leaders of each association to which they belonged were selected. The responses were: 1 = by outside agents; 2 = each leader chooses a successor; 3 = by small groups of members and 4 = by a vote of all members. The village index was computed by averaging the number of responses in the village.

The concept of "leader heterogeneity" is defined as the degree to which leaders of associations within a village differ from the rest of the people they lead, in terms of education and wealth status, computed as an additive index from responses to two questions concerning the wealth and educational status of each association leader. The responses were coded as follows: 1 = lower than most members; 2 = the same as most members; 3 = higher than most members. The village index was computed by averaging the number of responses in the village.

Participation in associations and group characteristics were measured following the work of Narayan (1997). Each respondent was presented with a list of different categories of associations compiled with the assistance of key informants and asked to indicate the associations in which household members participated. For each category of association in which the household had membership, the most active household member in that association was asked about the homogeneity of the association (in terms of religion, ethnicity, gender, education and income status), group size, number of meetings held over the last 12 months and member participation rate in those meetings, decision-making norms when selecting leaders or making other important decisions and leader heterogeneity (in terms of livelihood and education).

Narayan (1997) measured group characteristics using five attributes: (1) kin heterogeneity of membership; (2) income heterogeneity of membership; (3) group functioning; (4) group decision making; and (5) voluntary membership. The choice of these characteristics is based on the assumption that the contribution to social capital of being a member of each group is greater if the group is more heterogeneous across kinship groups, more inclusive, horizontal and better functioning. Next an index for village social capital was computed by combining the frequency of membership in associations with the characteristics of groups. Related measures have also been developed and applied in a different empirical context by, amongst others, Grootaert (1999) in Indonesia and Maluccio et al. (1999) in South Africa.

The advantage of a measure based on membership and group characteristics is that it incorporates both people's propensity to engage in collective action and the nature of the social interaction (whether bridging or bonding). However, by aggregating these two aspects of social capital into one index, it becomes difficult to know which aspect is more important when used in examining relationships. In the present study different aspects of associations are examined individually rather than combined in an index. It is important for policy makers to know which aspect is responsible for the observed relationship.

Bilateral transfers from social networks also constitute an exogenous income to the household. Bilateral transfers in the form of gifts, informal credit and labour are part of the major resources that rural households in developing countries exchange. These transfers can influence the demand for banana management practices directly when used in implementing the practices or indirectly by reducing risk aversion (Fafchamps and Lund, 2003). Bilateral transfers accruing to the household were measured for a period of 12 months by asking respondents whether the household had received any cash or any other items in the form of a gift, donation or free labour from people other than household members during the last 12 months. For items received by the household, the respondent gave the total amount received and for non-cash items the concept of willingness to pay was used to attach a value to the items so as to standardize those items across households. Similar questions were asked about the household expenditure on each item in the social network for the same period. The net transfer for each item was computed as the difference between receipts and expenditure within a social network (Table 15). The data allows for estimation of the effect of different forms of bilateral transfers on technology adoption. Since the majority of the transfers are not conditioned on the occurrence of a shock, the problem of selection bias is not expected to be important in the data. As such, the actual amount of transfers received is a representative measure of the household's access to bilateral transfers in the sample[4].

7.2.5. Independent variables used in social capital models

It is assumed that the decisions characterising participation in associations or private social networks specified in section 7.1. 3 are based on net benefits from participation. The net benefit that household (i) located in community (j) derives from participation

[4] Use of ex ante insurance networks in the analysis did not change the results.

in an association can be modelled as a function of individual-specific variables and community-level factors. The definition, measurement and hypothesized effect of each variable included in the estimation are described below.

7.2.5.1. Household-level factors

In the literature household-specific variables that may influence the net benefits from participation in an association or private social network include household wealth, education and demographic factors (such as age, gender, and household size). Household wealth (measured in terms of the possession of other household assets) may influence participation through budget constraints or the expected benefits (La Ferrara, 2002). The effect of wealth on participation in an association is likely to depend on the nature of the association. Because socially oriented associations are relatively cost-free, we expected them to be easily accessible by the poorer households and that wealth should not have much effect on participation. However, given the fact that the immediate need of most poor people is their survival, they are likely to see less benefit in socially oriented associations[5] while the richer households may derive benefits in the form of social standing in the village. We also expect wealth to positively influence participation in economically oriented associations because the benefits derived from these associations are high and because they require members to contribute resources that may not be affordable by the poorer households. Overall, wealthier households are expected to invest more in social capital than poorer households.

Education facilitates information acquisition about other people and hence increases the ability to cooperate as well as the confidence of the individual to speak up in a group. Less educated individuals may feel intimidated, especially when the group has better educated members with a high social status. In addition, education may enhance the productivity of social capital. Hence it is expected to be positively associated with participation in social capital accumulation. Age influences the way the individual

[5] It should be noted that specific associations are not mutually exclusive. Most social associations also provide informal economic services and functions, such as access to information that reduces the cost of transactions. Likewise, economic associations may provide opportunities for people of similar beliefs to interact. Therefore, the classification adopted here is based on the degree of orientation and is meant to simplify the exposition.

discounts the future and lowers the propensity to invest in social capital (Glaeser et al., 2001). Gender may create differences in participation because of differences in roles or constraints. Women, as compared to men, have a high opportunity cost of time, and gender norms in the community sometimes constrain their social interactions. Hence we expect male-headed households to invest more in social capital than female-headed households. Age, gender and education are personal characteristics of the household head. Other household demographic factors consist of the number of household members in different age categories. The descriptive statistics of these variables are summarized in Table 16.

Table 16. Household-level factors hypothesized to influence social capital

Variable	Variable description	Expected effect	Mean	SD
Age	Age of household head in years	-	49.87	15.41
Household members below 15	Number of household members below 15 years of age	+	2.53	2.41
Household members between 16- 50	Number of household members aged between 16 and 50 years	+/-	2.00	1.33
Household members above 50	Number of household members aged above 50 years of age	-	0.52	0.66
Gender	Gender of the household head = 1 if household head is male and = 0 if female	+/-	0.74	0.44
Education	Years of schooling of household head	+	5.58	4.35
Landholding in 2001	Total acres of land owned in 2001	+/-	3.79	3.28
Livestock in 2001	Total livestock units owned in 2001	+/-	1.15	2.14
Duration of residence in the village	Number of years the household has resided in the village	+	31.96	16.19
Distance to nearest post office	Distance in kilometres from the homestead to the nearest post office	+/-	1.42	2.39
Relatives	Number of relatives the household members can talk to freely and approach for help in case of a problem	+/-	4.67	5.78
Farm production orientation	A dummy = 1 if household head is primarily employed on the farm and = 0 if household head is primarily employed off the farm	+/-	0.66	0.47

Livestock units and the size of landholdings are the wealth indicators used in the analysis. Since decisions regarding the accumulation of assets and decisions to join associations may be made simultaneously, the previous household asset position in 2001 was used. A sub-sample of 100 households interviewed in the sample were also interviewed in 2001 under the project entitled “Reviving banana productivity in central Uganda” (see Chapter 6). Since 91 households were re-interviewed, the selection bias associated with attrition was not considered to be a problem. Using explanatory variables observed in 2001 in models of social capital measured in 2003-2004 helps reduce the endogeneity of the explanatory variables in the social capital formation. Of course, although observed in 2003-2004, some households could have joined these associations even before 2001 and the independent variables would still be endogenous. To check this, each respondent was asked to give the year he/she joined the association. For the sub-sample used in this analysis, the majority of the associational memberships (65 per cent) reported were acquired after 2001, thus supporting the assumption that the explanatory variables measured in 2001 are exogenous in the social capital equations.

The duration of residence in the community, measured as the number of years the household has lived continuously in the same village, proxies the household's connectedness to the community and hence its willingness to cooperate with others. The initial endowment of private social networks was measured as the number of relatives household members spoke to intimately and could rely on in times of need.

Distance from the homestead to the nearest post office is a proxy for the cost of communication. Households residing in less accessible locations may face high costs of communication that could constrain social interactions, since in rural areas face-to-face interaction is the main channel of social capital formation. This variable also captures the degree of physical access to markets and hence the interdependence among households. Hence the nature of association is not clear. Farm production orientation is a dummy variable measuring whether the household head is employed as a full-time farmer or works part-time off the farm. All household characteristics, except for wealth, were contained in the 2003-2004 data.

7.2.5.2. Village-level factors

Village-level factors included in the analysis are the social and economic heterogeneity of the village and the institutional environment, described in Table 17. The effect of social and economic heterogeneity is ambiguous. It can reduce group participation (Alesina and La Ferrara, 2000; Alesina and La Ferrara, 2002) or increase it if the population stratifies into homogeneous groups (Cornes and Sandlers, 1986). Social differences may imply differences in beliefs and social norms that could constrain free interaction in the village. Similarly, economic heterogeneity can create barriers to free interaction associated with differences in preferences that make it difficult for a group to reach an agreement.

Social heterogeneity was represented in the analysis by the continuous index measuring the degree of ethnic[6] fragmentation, computed with the formula used by La Ferrara (2002) as s:

$$F_j = 1 - \sum \Psi_{hj}^{\ 2};\ h = 1,\ldots h_j$$

where Ψ_{hj} is the share of respondents in village *j* who belong to the ethnic group *h*, and in each village there are h_j number of different ethnic groups. The index represents the probability that the two individuals drawn from the same village belong to different ethnic groups. Economic heterogeneity was represented by the village's educational heterogeneity, computed on the basis of the educational level of the household head as the standard deviation of the years of education of household heads in the village in 2001.

Table 17. Village-level factors hypothesized to influence social capital

Variable	Variable description	Expected effect	Mean	SD
Ethnic fragmentation	A continuous index representing the probability that two individuals drawn from a village in 2001 are from different ethnic groups	+/-	0.48	0.13
Education heterogeneity	This is computed as the standard deviation of the village's level of education	+/-	4.03	1.12
Number of NGOs	Number of NGOs operating in the village	+	1.61	0.79

[6] The concept of ethnicity is used here to refer to a social group of people with a shared tribal affiliation based on patrimonial lineage.

The number of NGOs active in the village represented the institutional environment. The institutional environment measures the incentives and level of information about group formation in the village and hence is expected to have a positive effect. Village asset heterogeneity variables used in the analysis were also contained in the 2001 data.

7.3. Summary

This chapter provides a link between the theoretical analyses presented in Chapter 5 and the econometric estimation procedure. A detailed overview of the economic relationships central to this study is presented and the estimation procedure for each economic relationship is indicated. Two important decisions regarding banana management are estimated in a two-step procedure, with the first step estimated as Probit and the second step as OLS, while correcting for selection bias where necessary. The first step aims at estimating the probability of using a given management practice and generates estimates that are used to compute the inverse Mills ratio included in the second step as one of the explanatory variables. The second step of the estimation focuses on the extent of use of management practices based on the relative demand for the practice. In the estimation of management decisions social capital is taken to be exogenous. The chapter also describes an empirical model used to estimate the determinants of social capital, hypothesized to be a function of household characteristics and village social and economic attributes. The decision to join an association is estimated as a Probit, while intensity of membership and private social networks are estimated as Poisson or Negative Binomial models.

Three methodological approaches to technology adoption studies have also been reviewed. Despite the inherent methodological limitations, cross-sectional studies still dominate research on technology adoption because they are simple and low-cost compared to panel data studies. The chapter ends with a detailed list and the respective measurement methods of the variables used in the analysis. Empirical methods and strategies were developed to measure technology adoption as well as social capital, borrowing ideas from the literature.