CHAPTER FIVE: RESEARCH PROCEDURES, METHODOLOGY AND TECHNIQUES

5.1 INTRODUCTION

Cross-sectional data on 153 farmers in the study areas were used. The data were collected by means of personal interviews in a sample survey in 1998. Appendix 1 shows the location of the survey areas. The next section of this chapter discusses the sampling methodology and techniques used in collecting the data. Section three presents a review of theories on the econometric models used in the study.

5.2 SAMPLING METHODOLOGY AND TECHNIQUES

The population from which the data for this study was collected is farmers in the Lowveld and Northern Regions of the Limpopo Province. Simple random and stratified sampling was employed to obtain the sample of small scale/emerging farmers for the study. Stratified sampling was employed because of the low number of borrowers in the study area. A random sample of regions in the province was done to select two regions as primary units. These were the Lowveld Region and the Northern Region. Lists of borrowers were obtained from financial institutions and other related organisations in the two regions, in addition to the list of farmers. At this stage, a stratified sampling was employed to select borrowers and non-borrowers for the study.

Three factors were considered in deciding on the size of the sample.

- The degree of precision required between the sample population and the general population[8].
- The variability of the population.

---

[8] *In deriving the number of farmers to be used for the study, 95% confidence level with a sample precision of ±0.05 was used. Past research done by Mokoena et al, (1997) indicates that 29% of small-scale farmers are borrowers.*

- The method of sampling.

Taking cost considerations and the above factors into account, a sample of 153 farmers was interviewed using a structured questionnaire. Of the 153 households sampled, 45 were borrowers and 108 were non-borrowers. The regional distribution is presented in Table 5.1.

Table 5.1: Distribution of borrowers and non-borrowers

Region	Borrowers	Non-borrowers	Total
Lowveld Region	25	65	90
Northern Region	20	43	63
Total	45	108	153

The information collected in the survey included data on household demographics, land tenure, agricultural production, livestock ownership, asset ownership, and credit and savings. The agricultural data covers the 1998/99 season. Control questions were included in the questionnaire to verify the consistency of the answers given by the respondents on various questions. In addition, the enumerators were instructed to use other control questions not included in the questionnaire whenever there seemed to be inconsistencies in a respondent's answers. A good deal of time was further spent in the field and in the office checking the consistency of answers to the questions.

5.3 MODEL SPECIFICATION

Two econometric frameworks are described in this section. The impact of credit on small-scale farmers in the Limpopo Province is estimated using an endogenous switching regression model patterned after the models of Maddala 1986 and Quandt (1973). Variants of the model have been used by Freeman *et al* (1998); Sail & Carter (1996) and Lapar *et al* (1995). The second econometric model is the logistic regression, which is used to locate and assess the factors limiting small-scale farmer's acess to institutional credit. Tabular analysis on the differential access to credit among small-scale farmers is also presented.

5.3.1 Switching regression: Measurement of the impact of credit and its shadow price

While credit interventions and credit market liberalisation policies have been justified on the grounds of lack of access to credit, the impact of improved credit access on agricultural production is only weakly understood because of identification problems which hamper the measurement and estimation of credit effects (Carter, 1989:13). Some studies (Taylor *et al,* 1986; Garg *et al*, 1977, among others) attempt to identify the effects of credit by estimating separate production functions for borrowers and non-borrowers and then proceed to compare the estimates. One of the weaknesses in this approach is the implicit assumption that all borrowers and non-borrowers are, respectively, homogenous with respect to their credit demand/supply situations. This assumption is often not valid, since many actually have sufficient liquidity from their own resources. These farmers may not need credit and it cannot automatically be assumed that they cannot obtain credit. Others cannot borrow because they are not creditworthy. Similarly, the marginal effect of credit may actually be zero for borrowers (Feder *et al*., 1990:1151).

David and Meyer (1980), pointed out that empirical problems may arise from the likely heterogeneity of credit recipients and non-recipients, and this can affect the true credit effect. This method cannot identify what proportion of the differences in productivity is due to differential credit access and what proportion is the result of other characteristics, which systematically differ between recipients and non recipients. The two groups potentially differ in terms of their:

1. observable characteristics such as endowments of land and market access;

2. unobservable characteristics such as endowments of farming; and entrepreneurial skills (Carter, 1989:15).

Carter (1989:16) pointed out that because of their potential differences in endowments, there are four competing (but not mutually exclusive) explanations of the descriptive statistical association between credit and farm productivity. The association may reflect a true credit effect if credit actually enhances the productivity of either observable or unobservable endowments. Alternatively, the association may reflect spurious correlation induced by the fact that credit recipients enjoy more favourable endowments of either observable or unobservable characteristics, and would exhibit the observed higher productivity even in the absence of credit. Fungibility contributes to the difficulty in obtaining a precise measure of the additional effect of credit.

While credit will in general enhance the opportunities of those who use it, it is very likely that credit recipients would still be performing better than non-recipients, even without credit, because the former may have better inherent characteristics than the latter. It is therefore very important to isolate the effect of credit on productivity from the effects of inherent characteristics of the farmer and farm (Lapar *et al*, 1995; Carter, 1989).

According to Lapar *et al* (1995:2) descriptive statistics often show differences in the average performance of borrowers and non-borrowers but do not measure the proportion of difference attributable to borrowing alone. In an effort to resolve this problem of attribute, some studies (Schultz, 1975, Taylor *et al*, 1986; Rana & Young, 1988) have estimated the effect of credit through single equation econometric models, which control for the influences of other productivity relevant variables which correlate with credit (Sial & Carter, 1996:776). According to Sial and Carter (1996:776), the reliability of such econometric efforts to resolve the attribution problem depends on the ability to measure and control all systematic differences between borrowing and non-borrowing farms and farmers. To them, one reason for questioning the adequacy of these efforts is the likely positive relationship between credit use and factors that are difficult to measure such as entrepreneurial ability, technical know-how and soil quality. They further argue that in the presence of latent factors correlated with credit, inference from a single equation - OLS approach would be subject to the same bias that confronts unconditional descriptive statistical analysis.

In addition to the identification problem, the potential heterogeneity of individual farmers raises further questions about the measurement of credit's effects. What is the appropriate measure of credit's impact on productivity – i.e. for what type of individual should the productivity effect of credit be defined? For an individual selected at random from the population at large? For an average individual who is a borrower? Or perhaps for the type of individual who is a non-borrower?

In other words, the heterogeneity of individual small-scale farmers implies that credit may have a different impact for the different kinds of small-scale farmers. Following Sial and Carter (1996) and Carter (1989), the following credit effect measures are used in the analysis:

- Random credit effect measure: It determines the effect that credit would have were it given to an individual selected at random from the overall population of small-scale farmers. The expected value of the latent attributes is zero for such an individual. This type of credit effect measure is appropriate for estimating the output supply effects of a generalisated loan programme, which managed to break the extant credit allocation regime and achieve a widespread credit allocation.

- Conditional or counterfactual credit effect: it compares the output anticipated by an individual under his/her actual credit status with the output level which would be anticipated for the same individual were he or she observed in their counterfactual state. In other words, the counterfactual credit effect indicates the impact of credit on the output of individuals who choose to be or not to be borrowers. Counterfactual measures are like before and after comparisons. For instance, it can be used to estimate the potential output of a non-borrower when credit is used. Likewise, it can estimate the expected output of a borrower under conditions of no credit use.

- Marginal credit effect: Evaluated at zero loan size, it gives the shadow price of capital in the absence of credit. A shadow price above $1+r$ (r is interest rate) implies that working capital is a binding constraint on small farm profit maximisation, and that non-borrowers are credit constrained to the extent that an additional unit of loan would result in more than a unit increase in output. The marginal credit effect evaluated at average loan size gives an indication of the optimality of loan size.

The econometric framework used to define and measure the impact of credit follows Freeman *et al* (1998); Sial & Carter (1996); Feder *et al* (1990) and Carter (1989). An econometric model that takes into account the non-random sorting of the sample between borrowers and non-borrowers is used to segregate the impact of credit from the impact of latent and observable characteristics of borrowers and non-borrowers. An endogenous switching regression model is used to structure and resolve underlying identification problems (Madalla, 1983:223-228). This approach is an improvement over the conventional use of OLS in estimating the outcome supply equation and deriving credit effects from the estimated coefficients. By correcting for selection bias, this econometric approach yields consistent and unbiased estimates of the parameters.

5.3.1.1 Development of econometric framework for measuring the impact of credit

Let the anticipated output supply (P) be defined as a function of loan size ''L'' and other characteristics. Thus, the analysis is built on the function P = f (L+ other characteristics).

The anticipated output values for individual "i" can be written according to one of the two production regimes.

$$P_i \quad = \quad P_{ic} = (\beta_c{}^i Z_i + \alpha l_i) + (V_{ic} + \varepsilon_{ic}), \text{ if individual 'i' receives loan}$$

$$P_{in} = (\beta_n{}^i Z_i) + (V_{in} + \varepsilon_{in}), \text{ otherwise} . \qquad (1)$$

In this switching regressions specification, the base regime, denoted with a subscript ''n'' applies when the individual does not receive a loan. The second regime, denoted with a subscript ''c'' applies when the individual receives a loan. The right hand side variables are partitioned into those which are observed and those which are not. The observable variables are Z_i and l_i. Where l_i is a quadratic expression of the loan amount L. The function αl_i gives the impact of loans on output supply and is a non- linear function of L_i which admits diminishing returns to L. The vector Z_l includes market conditions, prices and resources.

The parameters β_K (k = n,c) give the impact of the observable variables on output supply, and are allow to vary between the two regimes in (1) to allow for the possibility that relaxing financial constraints may permit an individual to earn large returns from a given market opportunity and level of fixed factors. The latent variables are divided into those known to individual 'i' (the V_{ik}) and those which are not (ε_{ik}). The terms V_{ik} (k = c,n) give the impact on output supply of intrinsic farm and farmer attributes such as farming skill and managerial ability. Without loss of generality, it can be assumed that this latent variable is scaled to have a zero mean for an individual selected at random from the overall small farm production of both borrowers and non-borrowers ε $(V_{ik}) = 0$. The ε_{ik} terms denote conventional, unanticipated random supply shocks, which are unknown to the individual at the time the production decision is made. It is assumed that $E(V_{ik}) = 0$.

Sorting borrowers and non- borrowers into the two regimes involves the decision of the individual to apply for a loan and the decision of the lender to make a loan. This implies two selection criteria functions: the individual decision, whether or not to apply for a loan, and the lender's decision, whether or not to grant the loan. The analysis of model which depends critically on whether the two decisions are independent or correlated, is whether or not the covariance of the error terms in the two criterion functions are zero. If the covariance is zero, implying independence, then the estimation of the parameters of the model is feasible and tractable. However, if the covariance is not zero, implying non-independence, which in the case of borrower and lender decisions is a realistic assumption, the estimation becomes more difficult because the expression of the expected values of the error terms get messy. To go around this messy situation, the bivariate probit method is used if information about the decision making process of the lenders is available.

In the analysis, because of the unavailability of information about lenders' decision-making processes, a second best approach is used. A single probit equation, which is an approximation of the two-probit equation is used. The single probit equation, which includes factors effecting the individual's side aw well as on the borrowers' side is used to infer lender behaviour. Thus, a self-selection approach is used to model credit allocation here, although admission of bank rationing criteria would not alter the analysis.

Credit status can be represented by the binary variable $L_{i,}$ which equals one if individual "i" has credit and equals zero otherwise. L_i can be modelled as a result of a latent credit access variable ļ which is scaled such that an individual becomes a borrower when ļ> 0. The ļ is the sum of orthogonal components:

One component systematically related to observable variables, and a second component known to the individual but which is unobserved by the econometrician.

$$ļ = \gamma^{i} x_i + \eta_I \quad (2)$$

Where X_I is a vector of variables which influence loan amount, γ is a vector of parameters and η_I is an error component reflecting random and latent factors which influence the loan amount. The observed loan amount, L_i is truncated at zero such that.

$$L_i = \begin{cases} 1, \text{ if ļ } \gamma^{i} x_i + \eta_I > 0 \text{ or } \eta_I > - \gamma^{i} x_i \\ 0, \text{ if otherwise} \end{cases} \quad (3)$$

Equation (3) defines the sorting of individuals into borrowers and non-borrowers as a probit process. The expected output supply conditional on the endogenous sample separation process and observable characteristics can be written as:

$$E(P_{ic}|L_I = 1) = \beta c^{i} Z_i + \alpha l_i + E(V_{ic}|Li = 1) \quad (4a)$$

$$E(P_{in}|L_i = 0) = \beta_n{}^{i} Z_i + E(V_{in}|Li = 0) \quad (4b)$$

Where the notation indicating conditioning on observable factor Z has been suppressed. The conditional expectations on the right hand sides of (4a) and (4b) can be rewritten as follows:

$$E(V_{ic}|D_i = 1 \quad = E(V_{ic}|\eta_I = - \gamma x_i \quad (5a)$$

$$E(V_{in}|D_i = 0 \quad = E(V_{in}|\eta_I = - \gamma x_i. \quad (5b)$$

The problem of intrinsic productivity differences between borrowers and non-borrowers can be clearly seen in (5). If latent productivity attributes are systematically related to credit status, then the conditional expectations in (5) will be zero. For example, an individual with greater skill is likely to realise larger output supply (via V_{ic}) as well have higher probability of obtaining credit under non-random sorting (via η_I) implying that $E(V_{ic}|L_i = 1) > 0$ in the borrowers subsample. OLS estimate of the output supply function under these circumstances, will yield inconsistent estimates of the structural parameters, attributing the direct output effects of latent individual skill to the observed loan amount with which it is correlated. However, OLS gives the best linear estimate of the (gross) output supply gap between non-randomly sorted borrowers and non-borrowers.

This gross output supply gap can be written as:

$$E(P_{ic}|\, V_{ic}) - E(P_{in}|\, V_{in}) = \delta^i Z_i + \alpha l_i + E(V_{ic}|\, Li = 1) - E(V_{in}|\, L_i = 1) \qquad (6)$$

where

$$\delta^i = (ß_c - ß_n)$$

The problem of non-random sorting which underlies the inconsistency of OLS fortunately suggests a resolution of the estimation problems. The Problematic correlation between the V_I and $\prod_I$ indicates that the latter in fact provides information on the latent variable V_i. By using that information to control for the latent characteristics, V_{ic} and V_{in}, the parameters of interest can be consistently estimate. By assuming that the error vector ($V_{ic,}$ $V_{in,}$ η_I) is multivariate normally distributed with zero expectations and positive definite covariance matrix, a full endogenous switching regressions system can be written as :

$$L_i = \quad 1, \text{ if } \eta_I > - \gamma^i x_i$$

$$0, \text{ if otherwise} \qquad (7a)$$

$$E(P_{ic}|L_i = 1) = ß_c{}^{i} Z_i + \alpha l_i + \rho_c \lambda_i{}^{c} \quad (7b)$$

$$E(P_{in}|L_i = 0) = ß_n{}^{i} Z_i + \rho_n \lambda_i{}^{n} \quad (7c)$$

where:

$\rho_c =$ $Cov(V_{ic}, \eta_I)/Var(\eta_I)$ and $\rho_n = Cov(V_{in}, \eta_I)/Var(\eta_I)$ are population regression coefficients relating to the V_{ic} and V_{in}, respectively.

$\lambda_i{}^{c} = \phi(¢_i)/\varphi\ (¢_i)$ and $\lambda_i{}^{n} = \phi(¢_i)/1\text{-}\varphi\ (¢_I)$ are the estimates of η_I given borrower type and $¢_i = \gamma^{i}x_i/Var\eta_I$; $\phi(.)/\varphi\ (.)$ are the standard normal density and cumulative distribution functions, respectively.

A two-steps Heckman procedure is used to estimate the parameters of this system Maximum likelihood methods can also be used. After using a first stage probit estimate of ¢ to contract $\lambda_i{}^{c}\ (¢_i)$ and a $\lambda_i{}^{n}\ (¢_i)$, consistent estimates of ß and δ^{i} may be obtained through separated OLS regressions of the two conditional output supply functions in (7a) and (7b)

Alternatively, the following expression can be utilised to define a single regression function easing hypothesis testing.

$$E(Pi) = E(Pi|Li = 1)\ Prob\ (Li = 1) + E(Pi|Li = 0)\ Prob\ (Li = 0) \quad (8)$$

Substituting from (7) above, equation (8) can be rewritten as:

$$E(P_i) = ß_c{}^{i}Z_i + \delta^{i}\ [\varphi(¢_i)Z_i] + \alpha\ [\varphi(¢_i)l_i] + (\rho_c - \rho_n)\ \phi(¢_i) \quad (9)$$

The parameters of (9) can be consistently estimated by doing OLS on the entire sample using first stage probit estimates of ¢ to contract $\varphi(¢_I)$ and $\phi(¢_i)$ for use as regressors (See Madalla, 1983). The specification in (9) allows for the estimation of direct credit effects parameters, the α and the indirect credit effect parameters, the δ^{i} and the $(\rho_c - \rho_n)$.

While the direct effect parameters give the increase in output supply due to the use of loans, the indirect credit effects represent the additional returns to observable and unobservable endowments when credit is used. If the loans do not enhance the return to other factors, i.e., both δ^i and (ρ_c - ρ_n) are equal to zero, then (9) reduces to the following equation:

$$E(Pi) = \beta_c{}^{i}Z_i + \alpha[\varphi(¢_i)l_i] \qquad (10)$$

Equation (10) is a restricted form of (9), where credit has only direct effects, but these effects still cannot be estimated using OLS given the non-random sorting of individuals between production regimes.

Figure 5.1 shows alternative relationships which may exist between credit, L, and output supply with credit, P_i . The dotted line in figure 5.1 has a slope equal to 1 + r, where r denotes the opportunity cost of capital. If the financial market inefficiency hypothesis is correct, then the slope of the output supply function should exceed 1 + r when L = 0, indicating that the additional output obtainable with an incremental increase in credit exceeds the full repayment cost of the loan.

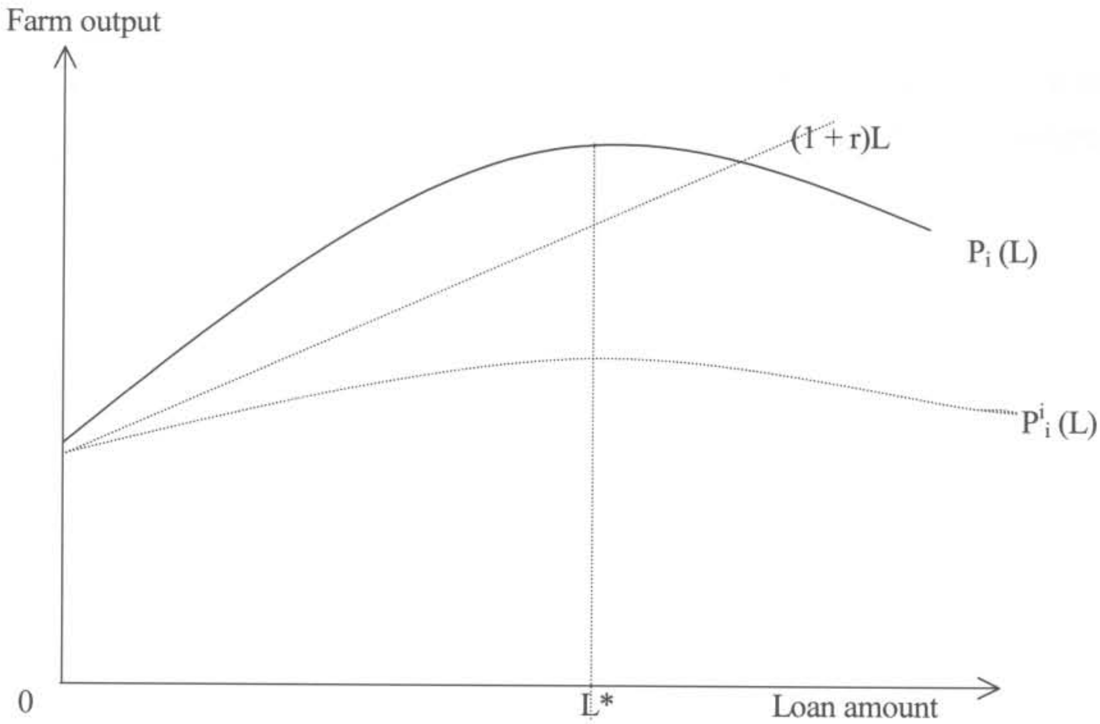

Figure 4. 1: Hypothetical impact of credit on farm output

The solid line labeled P_i (L) in the figure shows this possibility. The slope of the output supply function is equal to one plus the shadow price of the loan. Assuming diminishing returns to credit, there exist a loan size "L*" beyond which marginal returns fall below 1 + r. A loan larger than L* (used to purchase additional inputs) would decrease net income as each additional rand borrowed would produce less than 1 + r rands in increased output. Beyond loan size "L*", credit would no longer be a binding constraint on farm net income maximisation at interest rate r.

If the financial market inefficiency hypothesis is incorrect, and farmers are not general credit constrained, then one of two possibilities would describe the function shown in Figure 5.1. First, the output supply function would be completely flat if individuals who obtain credit do something with it other than investing it in agricultural production. Second, individuals could invest the loan in agricultural production despite obtaining marginal returns less than the full repayment cost of credit. While output would increase with loan size in this case, net farm income would decrease if loans were repaid in full. The dashed curve P_i (L) in the figure illustrates this latter possibility.

Definitions of credit measures

- Average or random credit effect-measures the impact credit would have on the productivity of a bundle of resources when farmed by an average individual selected at random from the overall population of individuals.

$$\begin{aligned} E(Pic\text{-}Pin) &= [\beta_c{}^i Z_i + \alpha l_I + E(V_i|L_i > 0)] - =[\beta_n{}^i Z_i + E(V_i|L_i < 0)] \\ &= \delta^i Z_i + \alpha l_i \end{aligned} \quad (11)$$

- Conditional or counterfactual credit effect - compares the output anticipated by an individual under the actual credit status with the output level that would be anticipated by the same individual in the counterfactual status.

$$\begin{aligned} &E(P_{ic}|L_i = 1) - E(P_{in}|L_i = 1) \\ &\quad = [\beta_c{}^i Z_i + \alpha l_I + E(V_{ic}|L_i = 1] - =[\beta_n{}^i Z_i + E(V_{in}|L_i = 1)] \end{aligned}$$

$$= \delta^{i} Z_i + (\rho_c - \rho_n)\lambda_i^{c} + \alpha l_i \text{ ....Borrower} \quad (12a)$$

$$E(P_{ic}|L_i = 0) - E(P_{in}|L_i = 0)$$
$$= [\beta_c^{\,i} Z_i + \alpha l_I + E(V_{ic}|L_i = 0] - = [\beta_n^{\,i} Z_i + E(V_{in}|L_i = 0)]$$
$$= \delta^{i} Z_i + (\rho_c - \rho_n)\lambda_i^{n} + \alpha l_I \text{ ....Non-borrower} \quad (12b)$$

- Marginal credit effect- Using equation (a) or (10) the partial derivative of output with respect to loan amount defines the marginal credit effect.

$$\partial E(P_i|L_i) / \partial L_i = \alpha_1 + 2\alpha_2 \, l_i \, [\varphi(¢_i)] \quad (13)$$

Evaluation of (13) at L = 0 gives the shadow price of capital in the absence of a credit programme A shadow prices above 1+ r implies that working capital is a binding constraint to small farm profit maximisation and hence provides a justification for credit or other intervention.

Evaluation of (13) at the average loan size renders some insight into the optimality of loan size under the existing credit programme. In addition, while the marginal effects of targeted credit speak to the efficiency of rural financial markets, measures of the total credit effects on output provide insight into the loan programme's economic impact.

5.3.1.2 Data specifications

The two stage switching regression model applied in this study uses a probit model in the first stage to determine the relationship between farmers' credit constrained condition and a number of socio-economic and credit variables. In the second stage, separate regression equations are used to model the production behaviour of groups of farmers, conditioned on a specified criterion function. Both the credit status and productivity could be influenced by different explanatory variables. The definitions of these variables and their hypothesized influences on credit status are presented in Table 5.2.

Table 5. 2: Data Specifications: Credit Status Equation (Probit)

VARIABLE	A PRIORI EXPECTATIONS
Dependent variable: **Credit Status** 1 = Access to loan, 0 = Otherwise.	-
INDEPENDENT VARIABLES	
Age of the household head in years	Age is posited to negatively affect the probability of being a borrower and hence, the loan size, in so far as older farmers may not be as active as younger ones in their farm activities.
Farm income (previous year)	Farm income is posited to negatively affect the probability of being a borrower and hence the loan size. Its sign is expected to be negative.
Non-farm income (previous years)	Non-Farm income is posited to negatively affect the probability of being a borrower and hence the loan size. Its sign is expected to be negative.
Financial assets (savings)	Financial assets (savings) - the sign for savings is indeterminate. It may either influence the lender to grant a loan to the borrower or will make borrowers use its their savings instead of a loan.
Remittances and pension	Remittances and pensions is posited to negatively affect the probability of being a borrower and hence the loan size. Its sign is expected to be negative.
Farm size in hectares	Farm size is posited to positively affect the amount of the loan as there is a greater need for variable cash inputs, and it is expected to increase capital access.
Family labour stock (members involved in farming)	Family labour stock and household size are hypothesised to reduce the amount borrowed because farms with large numbers of family workers might substitute labour for cash inputs.
Land ownership (Dummy: 1 if have title deeds, 0 if others)	Tenure - ownership, as opposed to rental and other forms of access to land is expected to increase the long run investment incentives and the collateral value of the land to lenders.
Regional Variables: (Dummy): 1 if Northern region 0 if Lowveld region	Its sign is indeterminate
Gender (Dummy: 1 if male, 0 if female)	Male are expected to have greater access to credit than female, hence its sign is expected to be positive
Education (Dummy variables) E1 E2 Level 1(no education to Standard 5) 1 0 Level 2 (Standards 6 – 10) 0 1 Level 3 (Above Standard 10) 0 0	Education- higher levels of education imply better technical know-how and farming skills, more information on markets and facilities provided by financial institutions. The coefficient is expected to be positive.
Repayment (Dummy Variable): R1 R2 i) Used credit, with good Record repayment 1 0 ii) Used credit with poor Record repayment 0 1 iii) Never use credit 0 0	Repayment record – It hypothesised that farmers who have repaid back their previous loans are usually perceived to be good clients and are provided with more credit in the following season. Hence, it will affect borrowing positively

Output Supply Equation

The dependent variable in the second stage regression is the log of total output value per hectare. All other continuous explanatory variables were expressed in logs. Expressing the dependent and continuous explanatory variables in logs provides dimension to measure the responsiveness of productivity to changes in input use. Since the coefficients of the regression equations are estimates of partial production elasticities, the larger the coefficient the higher the response of productivity to marginal changes in input use. Negative coefficients indicate that productivity actually declines as the level of input increases (Freeman *et al*, 1998:38). A quadratic form of the variable loan amount is included to account for the direct effect of credit on output (Sial & Carter, 1996; Lapar *et a*, 1995). The second stage regression did not include a farmers loan repayment records. The maintained hypothesis is that this variable is not likely to directly influenced output level. The definitions of these variables and their hypothesized influences on productivity are presented in Table 5.3.

5.4.2 Analysis of factors limiting small farmers' access to formal credit

The financial constraints on small farmers in applying modern technology optimally arises from their low level of income, as well as lack of savings. The only option left is to borrow from either the formal or informal credit sources or use meagre social benefits (such as pensions). Apart from the hypothesis that access to credit to small farmers is limited, it is also hypothesed that within the small farmer's group, farmers with bigger farms are relatively well off in terms of access to credit, thus, resulting in differential access to formal credit. In analysing differential access to credit among small farmers, the survey sample was divided into two sub-groups, i) operational holdings less than to 2 hectares and ii) those with operational holdings of 2 hectares or more. In addition, the whole sample was further divided into smaller sub-groups. The sub-groups are:

Group A:	≤ 1 ha.
Group B:	1.01 - 2 ha.
Group C:	2.01 - 3 ha
Group D:	3.01 - 4 ha
Group E:	> 4 ha.

Table 5. 3: Data specifications: Output supply equation

VARIABLE	A PRIORI EXPECTATIONS
Dependent variable: Output value	-
INDEPENDENT VARIABLES	
Age of the household head in years	Age – the number of years that a farmer has acted as a farm manager is expected to increase productivity.
Extension (Dummy): 1 if received extension services, o if otherwise	Extension contacts – extension agents in the rural areas have long been a strong arm in enforcing the adoption of innovations by farmers. All things being equal, extension contracts are expected to increase productivity.
Value of fertilizer used/ha (in rands)	Expenditure on variable inputs (seeds, fertilizer and other inputs) – It is hypothesised that farmers with relatively high expenditure on variable inputs are more likely to practice better management involving, among other things, the use of improved inputs.
Labour input/ha (in mandays)	Its sign is expected to be positive
Remittances and pension	Remittances and pensions is expected to affect productivity positively.
Amount of seeds used/ha (in rands)	This variable is expected to affect productivity positively.
Family labour stock (members involved in farming)	Family labour stock – the number of working age adults in the household represents fixed endowments, which will affect productivity positively.
Land ownership (Dummy: 1 if have title deeds, 0 if others)	Land ownership – possession of a legal title increases ownership security, and thereby increases the incentive to invest, which affects productivity positively.
Regional Variables: (Dummy): 1 if Northern region 0 if Lowveld region	Its sign is indeterminate
Gender (Dummy: 1 if male, 0 if female)	Its sign is indeterminate
Education (Dummy variables) E1 E2 Level 1(no education to Standard 5) 1 0 Level 2 (Standards 6 – 10) 0 1 Level 3 (Above Standard 10 0 0	Education – the number of years of formal schooling is an indicator of human capital, which positively affects efficiency.
Value of other inputs used	Input used is expected to affect productivity positively.
Amount of loan (in rands)	-
Loan amount square	-
PDC (probability density function or φ(C) in the model	-
Interaction terms of variables and CDF (cumulative density function or ϕ(C) in the model)	-

Following Amjad (1993), indicators of small farmers' access to formal credit, and factors affecting this were calculated in the following ways:

i) The proportion of formal credit received by a particular group to the total formal credit extended.

ii) By calculating access ratios, determined by the number or amount of loans extended to a particular group and the share of that group in the total sample or the total cultivated area. If the ratio equals one, the group has an average access to formal credit. Any number greater than one means above average access and if it is less than one, then below average access.

The ratios are calculated as follows:

$$Ratio\,1 = \frac{\text{Proportion of credit (number) received by the group to total formal credit}}{\text{Proportion of that group in the total sample households}}$$

$$\text{Ratio 2} = \frac{\text{Proportion of credit (amount) received by the group to total formal credit}}{\text{Proportion of area operated by the group to total area operated}}$$

iii) Logistic regression analysis was used to locate and assess the factors limiting small farmers' access to institutional credit.

According to Hair *et al.* (1998:276), logistic regression (logit model) and discriminant analysis are the appropriate statistical techniques when the dependent varaible is categorical (nominal or nonmetric) and the independent variables are metric. In this study logistic regression is used[9]. According to Hair *et al.* (1998:314), there are several reasons why logistic regression is an attractive alternative to discriminant analysis whenever the dependent variable has only two categories. Discriminant analysis is more appropriate when the dependent variable is nonmetric. However, when the dependent variable has only two groups, logistic regression may be preferred for several reasons.

9 *A similar type of analysis has been done by Adugna & Heidhues (2000) for Lume district in Ethiopia; Amjad (1993) for formal credit in North West Frontier Province in Pakistan; Anderson (1990) for Brazilian formal credit for small farmers and Sarap (1990) for rural Orissa in India.*

First, discriminant analysis relies on strictly meeting the assumption of multivariate normality and equal variance – covariance matrices across groups – assumptions that are not met in many situations. Logistic regression does not face these strict assumptions and is more robust when these assumptions are not met, making its application appropriate in many situations. Second, logistic regression can handle categorical independent variables easily, whereas in discriminant analysis the use of dummy variables creates problems with the variance/covariance equalities. Finally, logistic regression results parallel those of multiple regression in terms of their interpretation and the casewise diagnostic measures available for examining residuals.

The use or non-use of formal credit sources is therefore explained with the help of household characteristics using logistic regression analysis. In logistic regression one can directly estimate the probability of an event occurring. This analysis predicts whether an event will or will not occur and identifies the variables useful in making this prediction.

Accordingly, a farm household has either borrowed (Y = 1) or not (Y = 0) during the year in which the farm survey was carried out. The explanation of this binary variable requires the construction of a probability model that links it to a vector of factors, X (Greene, 1993). The probability of borrowing decision can then be expressed as:

$$\mathrm{Prob}(Y = 1) = F(\beta' X) \tag{1}$$

Where ß refers to the vector of parameters that reflect the impact of changes in X on the probability of borrowing decision. The choice of a particular form for the right hand side of the equation (1) leads to an empirical model. Adopting the logit analysis, the probability that a farm household makes a decision to borrow from formal sources is a regression model given by:

$$\mathrm{Prob}\,(Y=1) = \frac{e^{(\beta' X)}}{1 + e^{(\beta' X)}} \tag{2}$$

Using equation (2) the probability of borrowing decision could be written as:

$$\text{Prob}(Y=1) = \frac{1}{1 + e^{-(\beta' X)}} \qquad (3)$$

Equation (3) is a logistic cumulative distribution function where:

$$\beta' X = \beta_o + \Sigma\beta_i X_i = v_i \qquad (4)$$

Where:

$e =$ the natural logarithm

$\beta_o =$ the constant term

$\beta_I =$ the vector of coefficient

$X_i =$ the vector of explanatory variables, and

$V_i =$ the error term

The estimation of equation (4) using the maximum likelihood method helps to identify statistically significant explanatory variables. In the preceding discussions, a list of factors was identified that influenced accessibility of credit for small-scale farmers. Some of these factors are used in the analysis. It is hypothesised that borrowing from formal sources can depend upon total operated area; tenurial status; family labour; literacy status and age of the head of household; farm income; savings; awareness of formal institution, repayment records; and off-farm income. These characteristics are important in two ways:

a) they can influence the household demand for credit: and
b) potential lenders are likely to base their assessment of borrowers' creditworthiness on these characteristics.

It is difficult to completely separate the variables affecting either demand or access because decision making at both stages is based on almost similar considerations. Therefore, certain variables included in this regression are more related to small-scale farmers' demand for rather than access to formal credit. These include: age, farm income, sex of household head, and off-farm income. The data specifications are presented in Table 5.4.

Table 5. 4: Data Specifications: Credit Status Equation (logistic)

VARIABLE	A PRIORI EXPECTATIONS
Dependent variable: 1 = Access to loan, 0 = Otherwise.	
Independent Variables	
Age of the household head in years	Age of the head of the household is expected to have a negative effect, as comparatively young farmers are supposed to be more active in their farm activities.
Farm income (previous year)	The effect of farm income is indeterminate. A high farm income may reduce the demand for credit needed. On the other hand, it may make the farmer more creditworthy and in some cases, create a demand to expand production.
Non-farm income (previous years)	Non-farm income is assumed to reduce demand for credit and can be used to purchase cash inputs for production and/or even out consumption at times of need.
Financial assets (savings)	The sign for savings is indeterminate; it can influence the supply and demand size differently.
Remittances and pension	Remittances and pensions is assumed to reduce demand for credit
Farm size in hectares	Total area operated is expected to be positively related to access to formal credit because loans are advanced on the basis of ownership of land and bank officers generally expect that plenty of land will give plenty of output, enabling the loan to be repaid quite easily by the borrower.
Family labour stock (members involved in farming)	Family labour; its effect is indeterminate.
Land ownership (Dummy: 1 if have title deeds, 0 if others)	Ownership, as opposed to rental and other forms of access to land is expected to increase the long run investment incentives and the collateral value of the land to lenders.
Awareness of formal credit facilities in the area (Dummy: 1 if Yes, 0 if No)	Farmers' awareness of the formal credit channels available in their area is likely to have a strong bearing on their accessibility to formal credit. Ignorant farmers (of credit facilities) will implicitly squeeze themselves out of the formal credit market.
Gender (Dummy: 1 if male, 0 if female)	Male are expected to have greater access to credit than female, hence the sign is expected to be positive
Education (Dummy variables) E1 E2 Level 1(no education to Standard 6) 1 0 Level 2 (Standards 7 – 10) 0 1 Level 3 (Above Standard 10) 0 0	Literacy status can also influence farmers' access to formal credit institutions, and this is expected to be positive, because literate farmers are assumed to have better technical know-how and information about market and other facilities provided by the financial institutions.
Repayment (Dummy Variable): R1 R2) Used credit, with good Record repayment 1 0 ii) Used credit with poor Record repayment 0 1 iii) Never use credit 0 0	Whether a farmer is eligible for additional formal credit would depend on whether he has any outstanding loan that is overdue, and this variable would influence the demand for credit positively.

5.5 CONCLUDING REMARKS

The econometric frameworks discussed in this section will make it possible to analyse the impact and accessibility of credit on small-scale farmers in the Northern Province of South Africa. The four hypotheses postulated in chapter 1 will be tested with the results of the analyses presented in the next chapter.