Tanzania: A Hierarchical Cluster Analysis Approach

D. D. Ngaruko
Centre for Economics and Community Economic Development (CECED),
The Open University of Tanzania,
Email: deus.ngaruko@out.ac.tz

Abstract: *It is common for researchers and rural development policy stakeholders to describe smallholder farmers as a homogeneous group in terms of their demand for farm credit and farm investment behaviour. Given the diversity of factors such as farm credit products (input credit in cash, input credit in kind), farming systems (extensive Vs intensive farming, food crop Vs traditional cash crop production, crop production Vs livestock keeping), asset endowment, income sources and experience in farm credit borrowing, it is obvious that the demand for farm credit and use with which it is put are also diverse among farmers. Using survey data from Kibondo district, west Tanzania, we use hierarchical cluster analysis to classify borrower farmers according to their borrowing behaviour into four distinctive clusters. The appreciation of the existence of heterogeneous farmer clusters is vital in forging credit delivery policies that are not only appropriate for particular categories of farmers but also that do provide potential for reducing supply side transaction risks and costs.*

Key words: Smallholder farmers, hierarchical cluster analysis, farm credit supply,

INTRODUCTION

In Tanzania, like in some other agrarian economies dominated by smallholder farmers, the borrowers of farm credit as well as their households often are not homogeneous, even within the communities. Farmers in a given community have access to different resources. Some have more land, capital, or labour than others. In addition, knowledge and information are not owned and shared equally. Thus goals, resources and constraints differ between farmers. Soils and topography vary and seasons change. Because this variability influences what farmers can and wish to do, it is fundamental that active and potential farm credit suppliers understand real features of consumers of their services including farm credit. The failure by credit suppliers to recognize these differences between farmers has led to the mismatch between supply of credit products and liquid demand because they end up dealing with a smaller or possibly unrepresentative subset of potential borrowers (*Shreiner, 1997;* Conning et al, 2005). Alternatively farm credit suppliers tend to have very static view of farmers' resources and/or constraints. As a result, credit suppliers normally run the risks of developing credit product and credit supply contractual arrangements that are only relevant to a more restricted number of farmers than desired/expected (Dorward et al, 1998; Poulton et al, 2004). The first section covers the introduction, problem statement, research objectives and questions. The second section covers the conceptual framework of the hierarchical cluster analysis with highlights on methodology. Results are presented and discussed in section threewhereas the final conclusion of the paper is presented in section four followed by references and appendices.

Ngaruko Dominic Deus

Problem Statement

The understanding of real demand for farm credit by smallholder farmers, it is critical to identify and characterize clusters of farmers who share similar goals, resources and constraints in their socioeconomic and agroecological environments, because these farmers share similar farm input credit problems and thus require comparable solutions. However, the demand for financing agriculture based on various economic activities within the farm set-ups is not adequately researched and documented. These activities must be understood in order to adapt any farm credit services. This implies that the understanding of producers' strategies, their production and management choices according to their constraints, understanding how producers are integrated in the market and how this raises their farm financing difficulties requires scientific analysis of farmers' typology, their income and budget analysis as well as analysis of their farming paths. This study contributes to this missing knowledge on effective services delivery to the agriculture sector and more so to smallholder farmers as advocated in the current Tanzanian green revolution initiatives.

Study Objectives

The current study has three main objectives:
(a)   To explain an appropriate cluster analysis process of smallholder farmers in the study area.
(b)   To identify main categories of clusters of farm credit borrowers in the study area
(c)   To explain how farm business performance, features of farm credit and. demographic factors influence farmer's borrowing behaviour

Research questions

(a)   What are the factors to consider when clustering borrowers of farm credit?
(b)   Are demographic factors important in classifying borrower farmers?
(c)   Are the clusters of borrower farmers heterogeneous due to their differences in the access to land, labour and farming systems?
(d)   What are the farmer's business performance indicators?
(e)   What are the common features of farm credit?
(f)   What are the appropriate forms of farm credit for each cluster of borrower smallholder farmers?

CONCEPTUAL FRAMEWORK AND FARMERS CLUSTERING PROCESS

The process of identifying typology of borrower farmers was carried out in two complementing stages of cluster analysis (CA). The first stage of CA was a non-statistical participatory clustering commonly referred to as "Farmer's Own Classification of Farmers (FOCF)" which aimed at understanding farmers' view of characteristics of borrower farmers in their own institutional environment. FOCF was carried out during a two days stakeholders' workshop 1. The borrower farmers' survey was carried out to explore the most frequently noted characterisitcs during the FOCF exercise.

The second stage involved a Polythetic Agglomerative Hierarchical Cluster Analysis (PAHCA) to statistically confirm, redefine and characterize the categories elicited by farmers during FOCF exercise. The two clustering techniques are briefly explained in the following sections.

Farmers' Own Classification of Farmers (FOCF)

Farmers' Own Classification of Farmers (FOCF) of farm credit borrowers was done during the two days workshop comprising representatives of all farm credit stakeholders in the study area held prior to primary data collection. The goal of FOCF was to identify the important variables that farmers found relevant when classifying borrower farmers through participatory means. The rationale is that farmers have their own categories for classifying themselves. By eliciting these categories and their strengths and weaknesses, it enabled the researcher to shortlist the most important variables to be used in further statistical cluster analysis. The strength with this clustering technique is that it takes into consideration a totality of the behaviour of the borrowers. The technique too accounts for certain determinants of farm credit transactions such as guilt or shame of the borrower which are not statistically amenable. From the perspectives of Grandin (1988) FOCF can be described as the method based on the knowledge of local people who may be aware of assets and relationships that may not have been captured by survey data.

These include initiative, entrepreneurial ability, experience and social hierarchy or political relationships. Following FOCF exercise, the most important criteria used to classify borrower farmers were cited to be ownership and access to productive resources, farming systems, crop sales income, farm financing requirements and farm investment strategies. Survey was done to collect data on these variables which were key variables used in the hierarchical analysis.

Irrespective of the advantages, FOCF, like any other non-statistical cluster analysis methods is subjected to human biases. In addition, FOCF method cannot accurately come out with a clear-cut, rational number of independent clusters since the judgment of clustering by panellists is based on external qualitative behaviour of the borrower. The categories elicited from farmers may in some cases be self serving and value-laden. For example, it was not clear whether laziness, shamelessness, industriousness or willingness to form farmer groups refer to truly personal characteristics, describe a position within a social hierarchy or represent a value judgment by one group of farmers regarding others. In interpreting the data, all these issues were carefully observed in order to recognize the implicit value judgments and social relations present in the elicited categories.

Hierarchical Cluster Analysis

A hierarchical cluster analysis is used to reveal clusters within data set that would otherwise be apparent (Everitt, 1993). Hierarchical cluster analysis assumes that variables are independent of each other and that they must display normal distribution curve. Based on Ward's Method, a polythetic agglomerative hierarchical clustering

technique (PAHCA)[1] was used to classify farm credit borrowers into four distinct clusters. The summary of cluster analysis is presented in a dendrogram (Figure 1).

Choice of Variables used in the cluster analysis

Kydd (1982) points out that one of the most important stages of hierarchical cluster analysis is to select distinctive variables from the data set that are believed to be the most important. The variables included in the hierarchical cluster analysis must be of the same measurement type. The rationale for the choice of variables used in the current study is briefly explained below. The variables were chosen because of their impact on the decisions made by borrowers on whether or not to undertake farming business and/or to acquire external farm credit, as well as decisions related to farm credit repayments. The most important continuous variables included in the PAHCA were measures of access to productive resources (resource endowment), farming system and farm business performance.

Farmer's access to and use of productive resources

Demand for productive farm resources often exceeds supply hence a farmer facing resource scarcity allocates available resources to best advantage in an effort to achieve investment objectives as closely as possible. Farm credit is meant to reduce the relative scarcity of productive resources by increasing supply and/or improving the productivity of less productive resources. The productive resources considered in the current study are land, farm labour and capital. These resources are briefly described below.

Land ownership (ha) and farming systems

This was measured as the total land owned by the farmer, both productive and idle land in hectares (ha). If size of productive land is small, the farming system tends to be more intensified by using more own labour (or hired if labour constrained) or capital investment in terms of fertilizer application. Assuming that the farmer is constrained by lack of sufficient working capital to maximize land productivity, the farmer will be tempted to demand (outsource) more farm credit. Subsistence smallholder farmers tend to be highly self-sufficient and diversified whereas commercial smallholder farmers tend to specialize and concentrate on mainly one crop. Land abundant farmers can afford to increase own food supply and sometimes with some marketable surplus by practising crop rotations. Demand for fertilizer application to increase land productivity for these farmers is likely to be very low if more fertile land can cheaply be put into production (extensive production). Johnson (1990) argues that subsistence farmers tend to have horizontal diversification i.e. production of several commodities on same or

---

1Polythetic cluster analysis refers to clustering based on more than one characteristic (variable). Cluster analysis (CA) techniques can be divisive or agglomerative. A divisive method begins with all cases in one cluster. This cluster is gradually broken down into smaller and smaller clusters. Agglomerative techniques start with (usually) single member clusters. These are gradually fused until one large cluster is formed. Most CA techniques are hierarchical, i.e., the resultant classification has an increasing number of nested classes.

different farm/plot whereas commercial farmers tend to do vertical diversification i.e. many steps in producing a given product take place on same farm. For instance, a commercial maize farmer can retail maize on farm, undertake maize milling, pack the maize flour and deliver the flour to consumers under contract. The study area had relatively a low population density with slightly over 60% of land covered by fertile natural forest suitable for agriculture implying that there was no shortage of arable land. It's anticipated that farmers in the study area might be experiencing differences in access to land hence practising varying forms of farming systems, which in turn has an implication on their demand for farm credit such as fertilizer.

Farm labour (man-hours)
The borrower farmers were asked to mention number of hours they spent on farm activities out of 10 working hours per day. Kydd (1982) observed that whether on own farms or as casual farm workers subsistence and inexperienced smallholder farmers are expected to spend more time doing farm work than commercial or experienced smallholder farmers. Drawing from interviews with farmers it was obvious that farmers spending less labour in their farms hired farm labour which in turn increased their demand for cash credit (if outsourced) to meet wages for the farm labour. In some cases farm labour is supplied on in-kind credit where farm workers are paid at later time in form of either cash or in-kind such as exchange of farm labour for a pesticide application or for grain maize.

Capital (farm assets)
Many capital assets are used in farming. They include livestock, durable and consumable assets. Durable assets include buildings, fencing, soil and water conservation works, machinery and farm equipment. Consumables include livestock at hand, materials and crop output in store, and cash. The value of livestock and farm equipment were used as proxy variables for farm capital in this study. Livestock in many subsistence economies are used as store of wealth as well as medium of exchange in many transactions. Some farmers in the study area were exchanging live animals with farm input when faced with loan repayment difficulties. The high market value of farm equipment measures the value and quality of intermediate farm assets which can be liquidated or a lender can get hold of in case of failures to repay.

Farm business performance
Demand for farm credit is directly linked to the anticipated profitability from the farm or farming system to which the credit is to be used (Kelly, 2005). Whereas some forms of credit are needed to increase farm physical output (e.g. fertilizer, pesticide or improved seeds), other forms of farm credit (e.g. milling, storage or transport) are needed during post harvest activities to add market value to the crop which can increase farmer's negotiation power over the commodity's price. In order to capture overall impact of acquired credit (both before and after harvest) on farm performance, farm output was measured as net of crop sales income instead of physical output. Measures of farm performance used in the cluster analysis were average crop sales income (crop sales

income per hectare), monetary value of total farm credit borrowed and farm credit rate of return (total crop sales income per farm credit).

Total and average farm income (TShs/ha)

The average total income per hectare is a proxy measure of overall farm profitability. The average income was used to measure how best land was productive. A farmer who intensifies production by the application of fertilizers or improved agronomic and post harvest practices is likely to have higher average returns. Farmers who have intensified crop production are likely to have demanded more productive resources which imply more demand for farm credit. However, increase in farm output is not necessarily due to application of inputs to increase soil productivity but it could be due to extensive farming. In this case a farmer will need more labour and post harvest investment which all may tempt the farmer to outsource more of post harvest farm credit. Thus use of net crop sales per hectare is a rational performance measure that takes into account effect of farm sizes across farmers.

Farm credit rate of return (CRR)

Return on farm credit is measured as a ratio of net farm sales income to the value of total volume of acquired farm credit by the farmer. The ratio is computed for the total crop income as well as for each individual category of crops (i.e. traditional cash crops and food crops). The Farm credit rate of return (CRR) measures the total amount of Shillings generated for each Shilling of farm credit injected into crop production. The rate of return must at least be equal or greater than 1 for a viable farm performance. A ratio of 1 suggests that a farmer manages only to attain a break even point. It should be noted that in interpreting the ARR it is essential to ascertain for the volume of credit involved. Some farmers may have borrowed small amount of farm credit simply because they managed to use own stock of resources, or opted for extensive cultivation (for land abundant farmers) in which case the contribution of farm credit may not be very obvious. Nevertheless, for crops which require purchased input very few farmers could use own resources and as such supply of some input credit for certain crops such as cotton and tobacco were interlocked with commodity market.

Field Survey data

To overcome the weakness of FOCF method explained above a survey of 75 purposefully selected borrower farmers was subsequently done. The survey was conducted in three of the four wards of Kibondo district, western Tanzania. The district was also purposely selected to represent other geographically isolated districts in Tanzania. Data were collect using structured schedule of questions enumerated by trained village extension officers.

The survey included questions on variables such as landholding size, farming system, type and ownership of livestock and other assest, amount and form of farm credit obtained, farm input credit contractual arrangements performance of farm business and biodata.

Rescaled Distance Cluster Combine

```
C A S E    0    5    10   15   20   25
 Label   Num +---------+---------+---------+---------+---------+
Case 55   50
Case 56   51
Case 41   38
Case 45   42
Case 46   43
Case 68   60
Case 69   61
Case 3    3
Case 52   47
Case 14   13
Case 54   49
Case 53   48          CLUSTER 1
Case 12   12
Case 15   14
Case 51   46
Case 57   52
Case 58   53
Case 19   18
Case 18   17
Case 34   33
Case 42   39
Case 11   11
Case 35   34                              CATEGORY A
Case 67   59
Case 61   55
Case 20   19
Case 63   57
Case 37   36
Case 36   35
Case 8    8
Case 64   58          CLUSTER 2
Case 33   32
Case 43   40
Case 16   15
Case 2    2
Case 17   16                                                SAMPLE
Case 62   56
Case 31   30
Case 9    9
Case 73   65
Case 7    7
Case 28   27
Case 30   29
Case 74   66
Case 44   41
Case 32   31
Case 72   64
Case 23   22
Case 75   67          CLUSTER 3
Case 21   20
Case 22   21
Case 5    5
Case 26   25                              CATEGORY B
Case 49   44
Case 24   23
Case 25   24
Case 71   63
Case 1    1
Case 6    6
```

Ngaruko Dominic Deus

Case 27   26
Case 50   45
Case 10   10
Case 70   62   CLUSTER 4
Case 38   37
Case 60   54
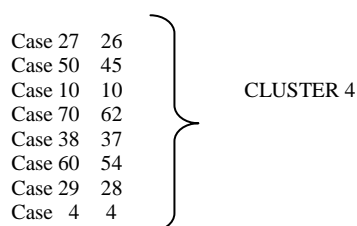Case 29   28
Case 4   4

Figure 1 Dendrogram of the PHCA for borrower farmers in the study area

Based on the information collected from borrowers themselves, it was possible to objectively cluster and characterize borrowers by use of the hierarchical cluster analysis.

Data analysis

Descriptive statistical analysis was carried out. The principal clusters of farmers generated through a Polythetic Agglomerative Hierarchical Cluster Analysis (PAHCA) are shown in the dendrogram (Figure 1). The statistical measures of central tendency were then used to further characterise the four clusters. The distinguishing features of these clusters and their associated statistics are presented in tabular forms. Test of significance between variables was done to using independent samples t-test.

RESULTS AND DISCUSSIONS

Clusters of farmers according to their access to productive resources and farming systems

The description of the characteristics of clusters according to their access to productive resources is based on table 1.

Table 1: Mean Values of productive resources and farming systems for clusters and sample

Characteristics		Mean Value				
		Cluster 1 N = 22	*Cluster 2* N = 18	*Cluster 3* N = 21	*Cluster 4* N = 14	*Sample* N = 75
Land (ha)	Total owned land (ha)	6.98	5.21	9.42	3.82	6.27
	Total land under crop production (ha)	2.93	4.53	6.86	5.97	5.31
	Area under cereal cultivation (ha)	1.05	3.57	2.03	2.36	2.24
	Aarea under beans cultivation (ha)	0.72	0.85	1.01	1.42	1.02
	Area under coffee cultivation (ha)	0.31	0.00	1.40	0.07	0.52
	Area under cotton cultivation (ha)	0.33	0.11	2.26	2.21	1.32
	Area under tobacco cultivation (ha)	0.32	0.00	1.09	1.04	0.65
	Total area under food crops cultivation (ha)	1.17	4.42	3.04	3.78	3.26
	Area under traditional cash crops (ha)	0.76	0.11	3.82	2.19	2.05
Labour (hrs per 10 hrs working day)	labour on own farm (hrs)	7.20	4.50	6.83	3.50	5.15
	hired on other farms (hrs)	1.50	1.50	0.00	0.00	2.50
	non farm employment (hrs)	1.30	4.00	3.17	6.50	2.35

Characteristics		Mean Value				
		Cluster 1 *N = 22*	*Cluster 2* *N = 18*	*Cluster 3* *N = 21*	*Cluster 4* *N = 14*	*Sample* *N = 75*
Capital Assets	Land utilization (used land/owned land) ratio	0.42	0.87	0.73	1.56	0.85
	Value of farm equipment (TShs)	50,909.0	67,188.8	49,590.4	267,728.5	94,920.0
	Market value of livestock (TShs)	170,782.7	124,132.7	445,504.7	536,100.0	304,701.4

Source: Survey data (2006)

Cluster 1:Land underutilization, subsistence farmers

Cluster 1 farmers have lower mean farm size than the sample mean. The mean farm size for this cluster is about 55% lower that of the sample mean. However the farmers seem to have excess land depicted by the fact that only less than half of land they own is put under crop production. This implies that the farmers can only utilize small portion of land they own. The farmers in this cluster allocated about 81% of total farmland to food crop production of which 53% is allocated for cereals whereas 28% was set for beans. The rest of the farmland (19%) was under traditional cash crop production. The farming system practised by cluster 1 farmers is likely to be small scale mixed farming where production of both food and traditional cash crops are jointly carried out. The mean value of farm equipment owned by the farmers in cluster 1 is also less than the sample mean by almost 46% or alternatively a quarter of mean value of farm equipment of cluster 4 farmers. The mean market value of livestock for cluster 1 farmers falls short of the sample mean by almost half. The main farm equipment is a hand hoe in numbers of 1-3 per farmer whereas the common types of livestock for cluster 1 farmers are small ruminants predominantly up to 10 goats per farmer. Cluster 1 farmers spend on average 5.20 hours out of 10 hours per working day in farming activities. This amount of farm labour is higher than the sample mean. Based on farm labour allocation, cluster 1 farmers can be described as full time farmers since they allocate about 72% of their daily labour on farming activities. In addition cluster 1 has the largest mean value of hours spent in farming activities than any other clusters.

Cluster 2: Land intensive, specialized food crop producers

Compared with other clusters, Cluster 2 farmers have the lowest mean farm size which is about 36% less than the sample mean. Farmers in cluster 2 allocate about 96% of their farmland to food crop production. Cereals account for 73% of total land equivalent to 128% of the sample mean. Very small proportion of total land (4%) is allocated to cotton as a traditional cash crop. Based on farming system, cluster 2 can best be described as farmers specialized in cereals (mainly maize) production. The cluster farmers have slightly higher mean value of farm equipment compared to those of clusters 1 and 3 but still lower than sample mean by almost 30%. Cluster 2 has mean value of livestock equivalent to only 40% of the sample mean. This is the smallest mean value of livestock compared to other cluster mean values. Cluster 2 has the mean farm labour of 4.50 hours out of 10 hours per working day which is lower than the sample mean. This suggests that cluster 2 farmers spend less than half of their daily labour on farm activities and thus could be described as part time farmers.

Cluster 3: Land abundant traditional cash crop and livestock producers

Cluster 3 farmers have mean farm size of about 18% above the sample mean. This implies that cluster 3 farmers on average possess larger farm sizes than those in clusters 1 and 2. Like cluster 1, cluster 3 farmers have mean farm sizes under food crops slightly below the sample mean. This cluster is predominantly composing of traditional cash crop producers. Compared to other clusters, cluster 3 has the least mean value of farm equipment. Cluster 3 farmers have mean value of livestock higher than the sample mean and over 300% that of cluster 2. The major type of livestock owned by cluster 3 farmers are numerous small ruminants such as goats and local breeds of dairy and beef cattle. Cluster 3 farmers have a mean number of hours they spend on farm activities slightly higher than the sample mean. This implies that like cluster 1, cluster 3 farmers are full time farmers spending about 68.3% of their daily working time on farming activities.

Cluster 4: Own land deficient, extensive single crop producers
With an exception of mean coffee farm size, cluster 4 farmers possess mean farm sizes above the sample mean for all other crops. Compared to other clusters, Cluster 4 farmers have the largest mean farm size equivalent to 36% higher than the sample mean. The farmers allocated about 63% of their total farmland to food crop production of which 58% was set for cereals while 23% was for the beans. Thus cluster 4 farmers can be best described as large scale (relative to sample means) practicing mixed farming. Both the mean values of farm equipment and livestock for cluster 4 are higher than the sample means and are the highest compared to clusters. The mean value of farm equipment owned by cluster 4 is about 182% higher than the sample mean whereas that of livestock is about 76% above the sample mean. Cluster 4 comprises the wealthiest farmers in terms of value of farm equipment and livestock. Cluster 4 spends on average the least time working on farms compared to other clusters. Like cluster 2 farmers, cluster 4 farmers spend less than half of their daily labour on farm activities, implying that the cluster comprises of part time farmers, probably using hired labour on their farms.

Clusters of farmers according to their farm business performance in relation to features of farm credit
Table 2 indicates that on average a hectare of farmland under traditional cash crops production yields about 56% higher returns than returns from a hectare of land under food crop production. This implies that it is more profitable to allocate more land to traditional cash crops than to food crops. However given the technocratic demand on traditional cash crops, very few farmers would wish to allocate more resources to traditional cash crops. In the following section we outline the four clusters of borrower farmers based on data on their farm business performance presented in Table 2.

Cluster 1: Farmers with the largest volume of farm credit, most efficient cotton farmers
Cluster 1 farmers borrowed on average, the largest amounts of both in-kind and cash farm credit. The mean value of total farm credit borrowed by cluster 1 farmers is higher than the sample mean by 64% and about twice as much as the mean value for cluster 4 farmers. On average cluster 1 farmers seem to have been charged a lower interest rate

than the sample mean interest rate on in-kind farm credit but a higher interest rate than sample mean interest rate on cash farm credit. This suggests that supplying cash farm credit to cluster 1 seems more risky than supplying them with in-kind farm credit thus reflecting higher interest rate for farm credit in form of cash. With an exception of mean average income from cotton and traditional cash crops as a whole which have mean incomes slightly above the sample mean, cluster 1 has mean values of average incomes for other crops below the sample mean. Cluster 1 has the highest mean value of average income from cotton compared to other clusters implying that cluster 1 farmer are the most efficient cotton growers. Farmers in this cluster have the least mean income per hectare from food crops.

Table 2: Mean values of features of farm credit and farm business performance for clusters and sample

Characteristics		Mean Value				
		Cluster 1 *N = 22*	*Cluster 2* *N = 18*	*Cluster 3* *N = 21*	*Cluster 4* *N = 14*	*Sample* *N = 75*
Volume of farm credit and its interest rate	total farm credit (TShs)	103,517.73	44,855.56	42,438.10	53,564.29	63,011.87
	value of in-kind farm credit (TShs)	43,872.27	16,244.44	4,133.33	25,371.43	22,661.20
	value of cash farm credit (TShs)	100,938.46	28,611.11	41,107.69	24,823.08	47,092.98
	interest rate on in-kind farm credit (%)	63.57	82.55	74.60	62.31	70.98
	interest rate on cash farm credit (%)	111.20	106.23	125.03	92.40	108.50
	Average Farm credit Interest Rate (AIR)	87.29	94.39	99.82	77.36	89.74
farm net income per hectare	maize farm income per ha (TShs/ha)	16,491.87	65,191.29	28,009.56	46,004.46	36,913.70
	bean farm income per ha (TShs/ha)	41,357.72	62,308.33	34,659.52	32,738.10	42,901.38
	cotton farm income per ha (TShs/ha)	138,417.50	0.00	74,864.42	60,708.33	82,463.96
	tobacco income per ha (TShs/ha)	70,000.00	0.00	63,386.22	159,596.67	78,075.40
	average income of food crop (TShs/ha)	21,749.76	89,900.10	27,141.92	53,740.86	45,587.32
	cash crop income per ha (TShs/ha)	72,140.91	0.00	60,738.93	117,380.74	70,903.73
	total farm income per ha (TShs/ha)	24,801.82	85,925.84	37,880.35	55,205.34	48,808.90
Farm credit rate of return (CRR)	CRR on food crops	9.38	13.59	15.41	16.27	13.36
	CRR on traditional cash crops	5.77	0.00	23.02	13.22	18.61
	CRR on total farm income	10.43	13.59	36.24	21.94	20.56
	percent full loan repaid timely	45.00	89.90	42.90	78.6	61.30
	percent failed to repay timely	50.50	11.10	57.10	21.40	37.30
	percent not due for repayment	4.50	0.00	0.00	0.00	1.30

Cluster 1 farmers have the least mean values of rate of return on farm credit for both food and traditional cash crops. The mean values are also less than the sample means. This implies that cluster 1 have the loan burden of borrowed farm credit. The farmers in this cluster can be regarded as inefficient borrowers which may affect their timely loan repayment. Less than half of cluster 1 farmers managed to repay loans in time. Their loan default rate was above the sample rate.

Cluster 2: Farmers with low demand for farm credit but with highest average income from food crops

The mean values of in-kind and cash credit for cluster 2 are lower than the sample means. Cluster 2 farmers are charged higher interest rate than the sample mean on in-kind farm credit but slightly lower interest rate than the sample mean on cash farm credit. Compared to cluster 1, cluster 2 farmers demanded less than half farm credit demanded by cluster 1 farmer. Being the cluster with the highest maize and bean average returns per hectare, cluster 2 exhibits the highest average total returns on food crops. The mean values of the average returns on food crops confirm earlier observation that cluster 2 farmers specialize in and thus optimize profit from food crop (mainly maize and beans) production. Cluster 2 farmers have the mean value of farm credit rate of return on food crops slightly above the sample mean but less than clusters 3 and 4 means. This is contrary to the fact that cluster 2 had the highest return per hectare. This observation implies that cluster 2 farmers used farm credit inputs more intensively compared to other clusters. Cluster 2 has the highest percentage of farmers who fully repaid the loans.

3.2.3 Cluster 3: Farmers with average demand for farm credit but with lowest average farm income

Like cluster 2 farmers, farmers in cluster 3 have obtained on average, below sample mean value of both in-kind and cash farm credit. Cluster 3 farmers have borrowed the smallest volume of farm credit. The mean value of in-kind farm credit for cluster 3 is about one fifth that of sample mean and about one tenth that of cluster 1. The mean interest rate on in-kind farm credit charged to cluster 3 is slightly above the sample mean. Cluster 3 farmers paid the highest mean interest on cash farm credit compared with other clusters. Cluster 3 has mean average returns for all crops below the sample mean. This is a reflection of the diseconomies of scale since cluster 3 farmers have the largest mean farm sizes. This suggests that with technology constraints, small farms can perform better than large farms. Compared to clusters 1 and 4 who produce traditional cash crops, cluster 3 have the lowest average returns on traditional cash crops. This is so irrespective of the largest proportion of total land allocated to these crops compared with other clusters. Cluster 3 farmers have mean values of farm credit rate of return on both food and traditional cash crops above the sample mean. As noted before, cluster 3 farmers borrowed smaller amounts of loans implying that in order to increase output farmers preferred extensive farming approaches (less demands for new technology) to intensive ones (higher demand on new technology) since the latter involves higher transaction costs. Cluster 3 has the highest farm credit rate of return on traditional cash crops. Comparing food to cash crops, cluster 3 farmers have higher farm credit rate of

return on cash crops than on food crops. Of all clusters, cluster 3 has the highest proportion (57.1%) of cluster members failing to repay the loan. This implies that cluster 3 is likely to be the leading loan defaulting cluster second to cluster 1.

3.2.4 Cluster 4: Farmers with very low demand for farm credit, highest average income from tobacco and food crops

The mean value of in-kind farm credit for cluster 4 is above the sample mean and higher than clusters 2 and 3 mean values. On average, cluster 4 borrowed the least volume of cash farm credit. The mean volume of cash farm credit borrowed by cluster 4 farmers was about half of the sample mean and about one fifth of cluster 1 mean.

Cluster 4 farmers paid the least mean interest rate on both in-kind and cash farm credit compared with other clusters. The mean interest rates charged to cluster 4 are smaller than the sample mean interest rates for both forms of farm credit. It can therefore be observed that cluster 4 farmers borrowed the smallest amount of farm credit and were charged the smallest interest rates compared to farmers in other clusters. The mean values of the average returns on beans and cotton for cluster 4 are below the sample mean. The rest of the crops for this cluster have average returns above the sample mean. Cluster 4 have the highest return on tobacco which has a mean value twice as much as the sample mean. The mean value of the total return from traditional cash crops for cluster 4 is about 66% above the sample mean.

Cluster 4 farmers are the most efficient tobacco producers. Unlike clusters 1 and 3, cluster 4 farmers are also efficient food producers second to cluster 2. Compared to cluster 2 which has average returns on food crops of about 97% above the sample mean, cluster 4 had mean value on food crops about 18% above the sample mean. Cluster 4 farmers have the highest mean value of farm credit rate of return on food crops.

Although the return per ha on traditional cash crops for cluster 4 was the highest of all other clusters, the mean farm credit rate of return on traditional cash crops for cluster 4 is lower than the sample mean. This is explained by the fact that farm credit is not sufficiently the only the necessary factor that may raise returns on traditional cash crops. Traditional cash crops press a high demand on technically sophisticated agronomic practices and farm labour which group 4 farmers may have failed to offer. The proportion of cluster 4 farmers who managed to repay the loans timely was very good and above the sample mean (78.6%) but not as high as that of cluster 2 farmers.

3.3 Demographic factors of farmers by clusters

Although not included in the hierarchical analysis, demographic variables were also very useful to characterize the four clusters of borrower farmers. Cross tabulations for these variables were worked out after the cluster case membership was established as shown on figure 1. The most important demographic variables were found to be age and education level of farmers.

Age groups by farmer clusters

Table 3: Cluster membership * age group of the farmer (years) Cross tabulation

Cluster		age group of the farmer (years)				Cluster Total
		below 30	31-40	41-50	above 50	
cluster 1	Count	0	5	11	6	22
	% within cluster membership	.0%	22.7%	50.0%	27.3%	100.0%
cluster 2	Count	10	8	0	0	18
	% within cluster membership	55.6%	44.4%	.0%	.0%	100.0%
cluster 3	Count	0	0	12	9	21
	% within cluster membership	.0%	.0%	57.1%	42.9%	100.0%
cluster 4	Count	0	9	3	2	14
	% within cluster membership	.0%	64.3%	21.4%	14.3%	100.0%
Total	Count	10	22	26	17	75
	% within cluster membership	13.3%	29.3%	34.7%	22.7%	100.0%

Cluster 1: Middle aged to elderly borrower farmers

Half of cluster 1 is composed of farmers aged 41-50 years with equal proportions of remaining half comprising of farmers aged 31-40 and over 50 years respectively. Cluster 1 does not contain any farmer from the youngest age category (i.e. below 30 year olds). Thus cluster 1 comprises of predominantly middle aged farmers who have witnessed and also have been affected by both pre and post economic restructuring policy regimes.

Cluster 2: Young, inexperienced borrower farmers

Cluster 2 comprises predominantly of less than 40 years old farmers with the youngest category (aged less than 30 years) comprising of 55.6%. It should be noted that all the sample farmers aged less than 30 years belong to cluster 2 thus no any other cluster has a cluster member aged below 30 years. These farmers are in the wealth accumulative stage whereby they use agricultural credit to increase their farm income which is invested in non farm income activities. They are dynamic but inexperienced farm credit borrowers. Thus they are interested in less risky short term investments in agriculture hence they are not involved in production of traditional cash crops which seem to be technically demanding and more risky than food crops.

Cluster 3: Elderly, most experienced borrower farmers

Cluster 3 contains the oldest categories of farmers. The cluster comprises of farmers aged over 40 years old with about 46% and 53% of sample farmers aged 41-50 years and over 50 years respectively. Although less dynamic, cluster 3 farmers are the most experienced traditional cash crop farmers and beneficiaries of pre reform government's agro subsidies. Agriculture to these farmers is taken as a tradition to nurture hence the only best alternative investment option. This implies that however unpromising incomes from farming may be these farmers will continue farming, at least using traditional farming practices.

Cluster 4: Middle aged farmers
Cluster 4 contains about 64.3% of farmers aged 30-40 years. This proportion is about 41% of sample farmers in that age category. The cluster composes of less than 12% of sample farmers aged over 40 years. Most of the 30-40 years olds are employed in non-farm activities and as observed before cluster 4 has the highest mean income from non-farm sources. Most of them are not natives of the district but are employees of the government and of other organizations working in the district. These farmers are less experienced in farming and farm credit borrowing than farmers in other clusters but they are knowledgeable of the repercussions behind farm credit borrowing. Most of them undertake farming as purely a commercial endeavour to supplement their non-farm incomes.

Education level of cluster farmers
Cluster 1 contains the largest proportion of farmers with no formal education compared to other clusters. Over 70% of cluster 1 farmers have at most primary education. Cluster 1 has some few farmers with post primary education. Over 90% of farmers in cluster 2 have attained at most primary education with 20.0% of them without any formal education.

Table 4: Cluster membership * highest educational level attained by the farmer Cross tabulation

Cluster	Highest Educational Level of the Farmer				Cluster Total
	No formal education	Primary education	Secondary education	Tertiary education	
Cluster 1	27.3%	45.5%	22.7%	4.5%	100.0%
Cluster 2	20.0%	73.3%	6.%	.0%	100.0%
Cluster 3	4.8%	76.2%	9.5%	9.5%	100.0%
Cluster 4	.0.0%	28.0%	64.9%	7.1%	100.0%
Sample Total	14.7%	48.7%	31.3%	5.3%	100.0%

Three quarters of farmers in cluster 3 have primary school education with about 19% of cluster farmers having at least secondary education. About 73% of farmers in cluster 4 have at least secondary education. This implies that cluster 4 contains the most elite cluster members of all clusters.

Independent samples t-tests for cluster variables
The statistical test was done to test the significance of the presence of differences between clusters described above. Independent samples t-test is used to test the null hypothesis that the sample variable means between clusters are the same. That is; $\bar{u}_1 - \bar{u}_2 = 0$ where $\bar{u}_1$ and $\bar{u}_2$ are mean values for clusters 1 and 2 respectively. An insignificant difference of variable means between two clusters confirms the null hypothesis implying that the respective clusters statistically exhibit similar behaviour for the particular variable in question. The t-test results for pairs of clusters for characteristics used to classify farmers are presented in appendices 1 - 5. To a greater extent, the results of

independent t-test conform to the description of the similarities and differences between clusters covered in the preceding sections.

CRITERIA FOR THE FOUR MAIN FARMER CLUSTERS
Figure 2 and Table 5 summarise main characteristics of the four clusters of smallholder farmers based on their farming systems, income, demand for farm credit, and return on farm credit.
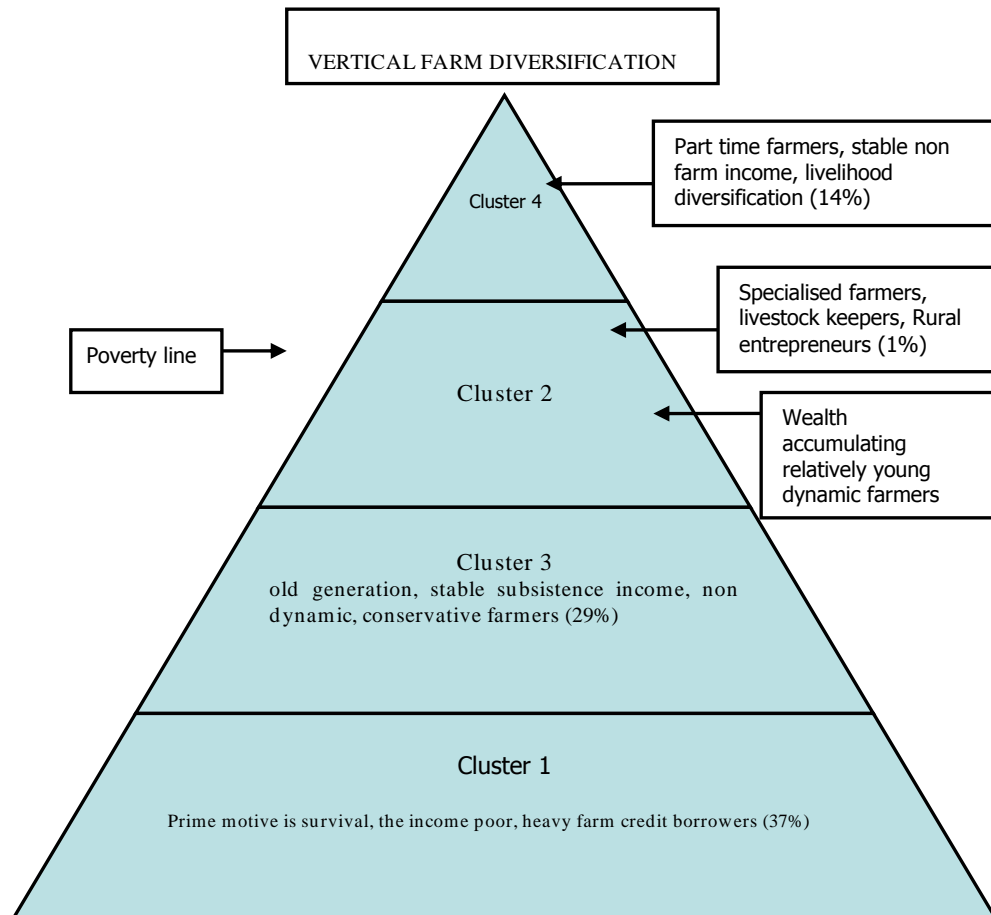

Figure 2: Pyramid showing proportion of farmer clusters relative to their participation in the current form credit arrangements

It can also further be concluded from Figure 2 that only a small proportion of farmers (about 14%) are likely to effectively participate in the current farm credit market. The top edge of Figure 2 shows proportion of farmers who are likely to be effective market participants whereas the bottom part indicates farmers who are unlikely effective participants in the current farm credit market arrangements. The higher the pyramid levels the more the likelihood that current farm credit supply arrangements are ideal for given cluster of farmers. Figure 2 suggests that cluster 4 farmers were effectively coping

with the current farm credit supply, followed by clusters 2, 3 and lastly cluster 1 at the base.

Table 5: Summary of characteristics of cluster of farm credit farmers

		Characteristics				
	Farming system	*Income*		*Demand for farm credit*	*Return on farm credit*	*high potential crop*
		farm	*non farm*			
Cluster1	Subsistence mixed crop farming	low	Low	high	low	cotton
Cluster2	Commercial Small scale	high	medium	medium	high	maize
Cluster3	Traditional Cash Crop farming	medium	low	medium	low	cotton, tobacco
Cluster4	Commercial Medium scale	high	high	Low	high	Tobacco, cotton, rice

RECOMMENDATION

Based on Table 5, Figure 2 and the cluster discussions covered in section 4, several hypotheses are recommended for further analysis to complement the findings from current study:

(a)   Well off and most successful farmers have lower demand for farm credit than the poor marginalized farmers

(b)   Traditional cash crops generate higher average returns on farm credit than food crops.

(c)   The well off farmers are charged lower interest rate on cash credit than the interest rate that is charged to relatively poorer farmers

(d)   Rate of return on farm credit is inversely proportional to volume of farm credit

(e)   It's difficult to understand the appropriate farm credit product due to heterogeneous farming systems as well as investment motives by different farmer clusters.

The proposed classification of smallholder farmers in the current study calls for the supplier of farm credit to consider smallholder farmers are very heterogeneous group in terms of their borrowing capability. The study has revealed that policies that categories farmers by their demand for credit may not work in some situations because borrowing seems to be a function of other farmer characteristics not considered by lenders as important. Non traditional factors to classify farmers in terms of the demand for farm credit include social-economic profile, ownership of high value crops like traditional cash crops, farming system adopted etc. Therefore there are difficulties in identifying ideal borrower due to differences in farming systems, hidden personal priorities, and willingness and capacity to repay the borrowed farm credit. The government and other development partners involved in sensitising formation of various sources of sustainable farm credit such as SACCOS or even linking farmers to credit sources such as banks

will find these findings very useful in forming more homogeneous groups of farm credit borrowers.

References

Conning, J. and C. Udry, 2005. Rural Financial Markets in Developing Countries. *Centre Discussion Paper No. 914*. New Haven, CT, Economic Growth Centre, Yale University.

Dorward, A., J. Kydd, et al., 1998. *Smallholder cash crop production under market liberalisation: A New Institutional Economics Perspective*. Wallingford, CAB International.

Everitt, S. B., 1993. *Cluster Analysis.* 3rd edition, New York, Halsted Press.

Grandin, B. E., 1988. *Wealth Ranking in Smallholder Communities*. London, IT Publications.

Johnson, S., K. Daniel, et al., 1997. The unofficial econoy in transaction. *Brookings papers on Economic Activities* **2**: 159-329.

Kelly, V., 2005. Farmers' Demand for Fertilizer in Sub-Saharan Africa. *World Bank Working Paper*(August 18).

Kydd, J. G., 1982. *Measuring Peasant Differentiation for Policy Purposes: A report on Cluster Analysis of the Population of the Lilongwe Land Development Programme for 1970 and 1979.* Lilongwe, Malawi.

Poulton, C., J. K. Ndufa, et al., 2004. *The viability of seasonal agricultural lending in Africa: Experiences from SCOBICS in Western Kenya*. Department of agricultural Sciences, imperial college (UK) and KEFRI research centre (Kenya).

Shreiner, M., 1997. *Rural financial markets in Argentina: Access to financial services and institutional performance*. Microfinance risk management, USA.

Appendices

Appendix 1: Test of mean equality between farmer clusters for features of farm credit

Variable	*Statistic*	*Test for equality of mean between clusters*					
		C1 & C2	C1 & C3	C1 & C4	C2 &C 3	C2 &C4	C3 &C4
total credit borrowed (TShs)	t statistic	0.672	0.687	0.569	0.106	-0.491	-0.455
	sig (2-tailed)	0.509	0.499	0.575	0.916	0.628	0.652
Monetary value of in-kind farm credit (TShs)	t statistic	0.692	1.010	0.458	1.750	-0.792	-2.298
	sig (2-tailed)	0.496	0.324	0.651	0.098	0.436	0.039
cash farm credit (TShs)	t statistic	0.907	0.722	0.954	0.503	0.361	0.655
	sig (2-tailed)	0.382	0.482	0.359	0.622	0.721	0.523
interest rate on in-kind farm credit (%)	t statistic	-1.077	-0.618	.064	0.609	1.289	0.770
	sig (2-tailed)	0.289	0.541	0.950	0.546	0.209	0.448
interest rate on cash farm credit (%)	t statistic	0.220	-0.495	0.744	-0.711	0.587	1.133
	sig (2-tailed)	*0.827*	*0.629*	*0.464*	*0.485*	*0.563*	*0.269*

* significant at 5%; ** significant at 1% *** significant at 0.1%

Appendix 2: Test of mean equality between farmer clusters for asset endowment

Variable	Statistic	**Test for equality of mean between clusters**					
		C1 & C2	C1 & C3	C1 & C4	C2 &C 3	C2 &C4	C3 &C4
value of farm equipment owned by the farmer (TShs)	t statistic	-0.720	0.090	-1.079	0.896	-0.996	-1.087
	sig (2-tailed)	0.485	0.929	0.300	0.381	0.337	0.297
Market value of all livestock (TShs)	t statistic	0.720	-2.594	-2.249	-3.164	-2.580	-0.502
	sig (2-tailed)	0.476	0.014*	0.039*	0.004**	0.021*	0.620
number of hours actually engaged in farming (hrs)	t statistic	4.212	0.547	2.671	-4.213	-0.421	2.418
	sig (2-tailed)	0.000***	0.587	0.130	0.000***	0.678	0.024*

* significant at 5%; ** significant at 1% *** significant at 0.1%

Appendix 3: Test of mean equality between farmer clusters for farm size

Variable	**Statistic**	**Test for equality of mean between clusters**					
		C1 & C2	C1 & C3	C1 & C4	C2 &C 3	C2 &C4	C3 &C4