INTERNATIONAL CONFERENCE OF AGRICULTURAL ECONOMISTS
ICAE
29th | Milan Italy 2015
UNIVERSITÀ DEGLI STUDI DI MILANO AUGUST 8 - 14
AGRICULTURE IN AN INTERCONNECTED WORLD

# Understanding the smallholder farmer in South Africa: Towards a sustainable livelihoods classification

*By Louw Pienaar[1] and Lulama N. Traub[2]*

*[1]Western Cape Department of Agriculture, BFAP, louwpp@elsenburg.com*
*[2]University of Stellenbosch, BFAP, ndibong2@gmail.com*

*Smallholder agriculture in South Africa has been identified as the vehicle through which the goals of poverty reduction and rural development can be achieved. To fulfill such potential, the need arise to understand diversity among these households to formulate effective policy interventions. This paper develops a farm typology of smallholder households according to their dominant livelihood strategies. Using multivariate statistics, Principle Component Analysis (PCA) and Cluster Analysis (CA), farming households are grouped in 7 distinct, homogeneous clusters. Results show the importance of social grants, specifically old age pensions and child support grants, in determining livelihood strategies of many smallholder farmers in the former homeland regions. Further evidence suggests that only a small number of households are able to market their produce. Essential characteristics of this group are the prevalent labor market attachment, higher usage of family labor and access to credit.*

IAAE

## 1. Introduction

The year 2014 has been marked as the year of the family farm and yet again the role of smallholder agriculture is called upon to help feed the estimated 9 billion people in the world by 2050 (Christiaensen et al., 2010). The debate on the specific role of this sector remains much debated. Should rural development strategies depend on smallholder farms to make a significant contribution to development, employment creation and poverty reduction? (Cousins, 2013; Larson et al., 2014). Indeed, South Africa's National Development Plan (NDP) has assigned smallholder agriculture, to drive development in rural areas and to improve the livelihoods of at least 370 000 people, specifically in the separate-developed areas known as the former homelands (NPC, 2011). The ANC-led government has committed itself to expand the number of smallholder producers selling their produce from 200 000 to 250 000 by 2014, and to 500 000 smallholders towards 2020 (Aliber & Hall, 2012). These attitudes towards supporting smallholders are clearly illustrated by the increased budgetary allocation by the Department of Agriculture, Forestry and Fisheries in recent years, with R2.38 billion being allocated to smallholder support programs in 2014 (Aliber & Hall, 2012; DAFF, 2014).

It is argued that this support to smallholders should be based on a concurrent strategy which should co-exist in achieving both scale and impact amongst smallholders. The former on the basis of welfare and food security, supporting a substantial proportion of subsistence producers to increase production, while the latter to be focused on narrow empowerment towards commercialization of a few (Aliber & Hall, 2012). To do this it is essential to address two shortcomings that exist within the development of this sector (Cousins, 2013). First, detailed analysis is needed to understand the livelihood strategies which results from divesification amongs this hererogenous group of approximately 2.6 million farming households. Second, confusion exist by what exactly is meant by a smallholder or small-scale farmer in South Africa (Kirsten & Van Zyl, 1998; Ortmann & King, 2006). Using the term "smallholder" generically suggests that these farmers are relatively homogenous which tends to obscure class-based

differences between them and causes misleading assumptions about common interests in development planning (Cousins, 2010; Tittonell, et al., 2010). Together, these two issues make it very difficult to determine the nature of support to be given to farmers, as evidence suggest that previous programs have been ineffective to stimulate rural growth and poverty alleviation (Perret et al., 2005; Ortmann & King, 2006; Hall & Aliber, 2010; Aliber & Hart, 2009). These considerations are crucial and needs to be addressed if a smallholder-focused approach is utilized in order to design and implement effective rural development policies (Laurent et al., 1999; Cousins, 2013).

One way of addressing such challenges is the use rural livelihood analysis that comes in the form of farm typologies (Capillon, 1993). Typologies have often been used in the literature in order to understand structural changes in farming with regards to output, employment, farming intensity and impacts of policy reforms (Iraizoz et al., 2007). The approach used to develop the typology comes from the sustainable livelihoods (SL) approach found in the literature (Ellis, 1998; Ashley & Carney, 1999; Scoones, 2009). Multi-disciplinary in nature, the SL approach seeks characterize rural livelihood diversification amongst rural households which include agricultural activities, but also in economic, social, environmental and political perspectives (Scoones, 2009). Within the framework of rural development, a typology is a procedure (qualitative and/or quantitative) for developing and describing relatively homogenous groups of households with similar characteristics and needs, and which are expected to respond to external influences in a similar fashion (Perret & Kirsten, 2000; Tefera, et al., 2004).

This paper will develop a typology that will classify smallholder households in the former homeland areas of South Africa according to their dominant livelihood strategies (Ashley & Carney, 1999). Using one of the very few nationally representative datasets with detailed information on smallholder livelihoods, the Living Conditions Survey (LCS), distinct cluster groups are identified for further analysis. The literature on smallholder typologies in South Africa's rural areas are mostly done by using smaller case study areas and are mostly participatory in nature, using qualitative approaches (Laurent et al., 1999; Perret & Kirsten, 2000; Perret et al., 2005). In contrast, this typology will include the entire former homeland areas

of South Africa as it was demarcated according to the Apartheid legislation and will follow the quantitative approach using multivariate statistics, Principle Component Analysis (PCA) and Cluster Analysis (CA).

## 2. Smallholder Diversity in South Africa

The agricultural sector in South Africa is dualistic (Vink & Kirsten, 2003). It consists of a well-integrated, highly capitalized commercial sector with approximately 35 000 white farmers, producing around 95% of agricultural output on 87% of total agricultural land (Aliber & Hart, 2009). In contrast, the smallholder sector consists of around 4 million black farmers farming in the former homeland areas on 13% of agricultural land of South Africa (Aliber & Hart, 2009). This dualistic nature and division between the commercial, large-scale farming sector and the comparatively low productive, struggling smallholder sector is a direct result of historical patterns of dispossession and impoverishment, which systematically eroded historically successful land-based production systems and livelihoods in South Africa (Neves et al., 2009). Even after 20 years of democracy, this sector continues to be characterized by inequality in terms of the distribution of economic assets, support services, market access, infrastructure and income (Oettle et al., 1998).

The general characteristics and history of the smallholder sector development is well-documented in literature (Bundy, 1979; Essa & Nieuwoudt, 2003; Kirsten et al., 1998; Lahiff & Cousins, 2005; Thompson, 2000; Hamilton et al., 2012; Aliber & Hall, 2012). Today, the smallholder sector is known for its small farms that are labour-intensive, uses traditional production techniques, and often lack institutional capacity and support (Louw, 2013). Furthermore, the majority of these farmers are located in the former homeland areas as it was demarcated according to the 1913 and 1936 Native Land Acts, and production is mostly aimed at securing food to the household (Groenewald & Nieuwoudt, 2003; Lahiff & Cousins, 2005). Figure 1 shows the number of smallholder households in South Africa and it is clear that 80% of these farm extra source of food to the household. On the other hand, only a very small number of these farmers, approximately 160 000, farm to market their produce (GHS, 2013).

<Figure 1 here>

According to Fenyes and Meyer (2003) the majority of these rural farming households consist of women, children and elderly people. These rural households develop different livelihood strategies said to be conducive to the given opportunities and constraints in their specific environment (Tittonell et al., 2010). Ellis (1998) suggests that the phenomenon of livelihood diversification among rural households is as a process by which a diverse portfolio of activities and social support capabilities are chosen, voluntarily or involuntarily, as a response to a crisis, for improved living standards. More often than not, farming activities form an important part of these strategies as the majority of rural people are either directly or indirectly linked to agriculture (Pauw, 2007). Livelihood options amongst these households would either/or be farm-related, off farm (wage employment on other farms) or non-farm (non-agricultural wage employment and transfers) (Ellis, 1998; Perret et al., 2005). Thus, farming households in South Africa's rural areas typically pursue different livelihood startegies on the basis of the available natural, physical, human and financial capital available to them and these are also to a large extent dependent on biophysical and socio-economic conditions.

The general characteristics of smallholder households in South Africa are given Table 1. The sample is limited to farming households located within the former homeland regions. The average age of the household head is 56 year, with an average of only 5 years of education; the equivalent of grade 5. Indeed, the majority (55%) of these households were headed by females and the average household consisted of 5 individuals. In terms of income, the main sources of income is social welfare grants from the government, specifically old age grants and child support grants, with monthly incomes of R522 and R282 respectively. The mean salary income was R939 per month, although only a limited number of households were connected to the labor market. Remittance transfers from family members working elsewhere and does not live permanently with the family amounts to R123 per month. The mean total income for smallholder households were R2732 per month

<Table 1 here>

Focussing now on the production systems of the household, the average size of land utilized for farming is 0.86 hectares and approximately 2 family members supply farm labor. Unfortunately this survey does enumerate the intensity of such farm labor. In terms of animal numbers, the average number of cattle stock is 5, 5 sheep or/and goats and 8 chickens. Expenditure on farm

inputs is approximately R36 per month, indicative of the relatively small production systems being managed and reluctance or inability to invest in their farming systems to boost production. It should however be noted that many smallholders in these areas receive support either from government (400 000) or non-governmental support (53 000) with which they attain additional inputs (GHS, 2013).

Finally, it is important to understand the definitions used when referring to smallholders in South Africa. Ambiguous usage of descriptive word such as “small”, “small-scale”, “family”, “subsistence”, “emerging”, and even the word “smallholder” has been used to refer to the group of farmers included in this research. Globally, one of the more general approaches used to define smallholder farmers would be to assess the common characteristics of these farmers such as their land and capital access, exposure to risk and input technologies and market orientation (Chamberlin, 2008). However, in South Africa, the term “small-scale” is generally used to refer to the total number of farmers or households participating in any agricultural production. According to the National Department of Agriculture, Forestry and Fisheries (2012), this broader group of small-scale producers should be subdivided into two groups; Emerging farmers and Smallholder farmers. The former connotes to those approximately 200 000 farmers selling their produce, while the latter refers to all of the rest which produce for household consumption.

This paper proceeds in the development of an appropriate classification system of these farming households. The aim is to identify homogenous farming households which inherently choose various livelihood strategies in order to obtain better living standards. Livelihood diversification amongst smallholder households are often based on a variety of factors and condition which makes a classification on single indicator problematic. The analysis which follows introduces an appropriate classification method and based on these considerations a typology is developed to better understand smallholder farm household in South Africa.

## 3. Data, Methodology and Descriptive Statistics

### *3.1 Conceptual framework*

Agricultural systems comprise of a basic production unit, i.e. the farm, which has its own distinctive limitations and constraints and faces a heterogeneous decision-making environment. Given this diversity within agricultural systems, various schemes of classification have been developed and evolved over time (van der Ploeg et al., 2009). The essential steps in developing farm typologies are well documented in the literature and have followed one of, or a combination of, two main approaches: qualitative systems and quantitative systems (Iraizoz et al., 2007; Laoubi & Yamoa, 2009; Righia et al., 2011). The quantitative system approach is used to develop the typology of farm households in this analysis because of its strength in objectively identifying groups based on probability theory (Van Averbeke & Mohamed, 2006). In recent years many studies have utilized the quantitative approach in order to create farm typologies (Ballas et al., 2003; Emtage et al., 2005; Bidogeza et al., 2009; Laoubi & Yamoa, 2009; Gelasakis et al., 2012). Figure 3 shows the combination of pathways used to create farm typologies and the following sections will briefly explain the steps followed to develop the typology of farming households

<Figure 2 here>

Step 1 involves the selection of the theoretical framework on which the classification will be based on. According to Emtage (2004), a variety of different theories have been used as a framework to develop farm typologies which include farming styles (Van der Ploeg, 1990; Van der Ploeg, 1993; Van der Ploeg, 1994; Vanclay et al., 2006), sustainable livelihoods (Belsky, 1984; Perret & Kirsten, 2000; Dorward, 2002; Tefera et al., 2004; Perret et al., 2005; Babulo et al., 2008; Tittonell et al., 2010; Righia et al., 2011), farming context (Kaine & Lees, 1994) and market structure (Barr, 1996). The theoretical framework adopted here is the sustainable livelihoods (SL) framework. SL is a multi-disciplinary approach which seeks to not only look at agriculture, but includes economic, social, environmental and political perspectives (Scoones, 2009). According to Ashley and Carney (1999) SL approaches are based on the ever-changing

thinking about poverty reduction, the way poor people live their lives and the importance of structural and institutional issues.

The premise of sustainable livelihoods is that the effectiveness of development undertakings can be improved by 1) systematic analysis on poverty and its causes; 2) a better informed understanding of the opportunities for development and its impact on livelihoods and 3) placing people and the priorities they define as the central part of the analysis (Ashley & Carney, 1999). This framework is chosen due to its strong association with rural development research and its strength in describing diversity at a community level. Furthermore, In order to understand diversification among farming, SL allows for dynamic analysis of the different strategies farming households undertake to attain a higher standard of living. This is done looking at various characteristics that would define specific farming systems and includes income and expenditure, household characteristics, production activities and socio-economic indicators. SL recognizes that a specific livelihood encompass more than just income (cash and in kind), but includes social institutions (kin, family, village etc.), gender relations, property rights and a few others which would influence the strategies adopted by rural households (Ellis, 1998). This approach recognizes the importance of the household as the decision-making unit which base decisions on the households' available resources, objectives, personal and socio-economic views and the rules and norms of institutions that govern the use of resources available to the household (Emtage, 2004). It has been used extensively in the development of farm typologies (Belsky, 1984; Perret & Kirsten, 2000; Dorward, 2002 Perret et al., 2005) in the past and the outcome of the SL approach is designed to improve the livelihoods of poor households by improving their levels of well-being, food security, income and biophysical environment (Emtage, 2004).

*3.2 Variable Selection*

The variable selection is one the most important steps in the analysis and will have the greatest impact on the ultimate results (Kobrich et al., 2003). Based on the selected SL framework, variables are selected to best describe differentiation between households in terms of livelihood strategies. According to Kobrich et al. (2003) variables selection is generally based on three grounds; 1) the researchers' experience and knowledge of the study area, 2) the objectives of the

specific typology and 3) the quantitative information that is available. Furthermore, the selection of variables is also guided by previous typology studies for the appropriate usage in PCA. This implies that variables should be included which are sufficiently dependent on one another and, at the same time, not too strongly correlated. To do this, the Kaiser-Meyer-Olkim (KMO) (1970) measure and Bartlett's test of sphericity is used. This test is used see if the selected variables are appropriate and valid for PCA (Field, 2009; Bidogeza et al., 2009). The variables included in the development of the typology are given in Table 1 and are divided into the following categories: household characteristics, income and expenditure, production orientation and socio-economic status. The inclusion of the variables will be discussed below under each category.

3.2.1 Households Characteristics:

In livelihood analysis, the household demographics play an important part in understanding diversity and have routinely been included in typology research (Perret et al., 2005; Bidogeza et al., 2009). Social relations such gender, class differences and other social differences is central to livelihood analysis as these inevitably govern the distribution of income, work patterns, consumption and accumulation dynamics and will therefore be included in the typology (Scoones, 2009). Other important variables include the age, years of education, gender and marital status of the household head. Also included in this category is a variable giving the years of education of the most educated individual in the home. Furthermore, the household size and the number of family members for each of the following sub-groups: children, employed, elderly and the number of spouses living outside of household is included in the analysis. These variables matter as a livelihood encompasses social and kinship networks for facilitating diverse income portfolios, as well as gender relations (Ellis, 1998).

3.2.2 Income and Expenditure

Household income is naturally an important determinant of livelihood diversification. The main income sources used in the analysis include salary income, old age grants, child grants and remittance payments. Total income is also included as some auxiliary incomes; together with the main income sources indicate the family's financial position, while per capita food expenditure

gives an indication of the expenditure differences in food consumption. All of these values were included in current monthly Rand terms.

3.3.3 Production Orientation

The variables included to measure the households' farming orientation are limited to the questions asked in the survey. Productions variables include the number of family members working on the farm, hectares of land used for farming, stock numbers for cattle, sheep and goats, and chicken and the monthly expenditure on inputs by the household.

3.3.4 Socio-economic Status

The variables included in this section highlights key constraints and options households face with regards to their socio-economic and welfare outcomes (Ellis, 1998). Individual measures for food security, access to credit and asset status are created to be included in the analysis. Using standard PCA, the first component is often used as an index by using the un-rotated factor solutions from the analysis. This technique is commonly used to create indicators for food security (Demeke & Zeller, 2009; Nyaga & Doppler, 2009; WFP, 2009; Wineman, 2014), asset status (Sahn & Stifel, 2000; Booysen, 2002; Filmer & Pritchett, 2001; Booysen et al., 2008; Vyas & Kumaranayake, 2006) and exposure to debt (Motsoari, 2012; Chichaibelu et al., 2012; Anderloni et al., 2011). All three indices from the PCA are transformed using the min-max normalisation which transforms the indices so that the values are scaled between zero and one in order to ease interpretation (Visalakshi & Thangavel, 2009). Table 2 shows the variables used to compute the respective indices for food security, asset status and credit access.

<Table 2 here>

*3.3 Study Area and Data*

The study area is located in South Africa's former Bantustan homeland areas as mentioned before. The data used in the analysis comes from Statistics South Africa's Living Conditions Survey (LCS), which is a nationally representative survey conducted during 2008 and 2009 (LCS, 2008/09). This survey is unique in the sense that it is one of the only nationally represented surveys that captures details on smallholder production in conjunction with

important information on household income and expenditure, household characteristics and welfare outcomes. Surveys that do capture smallholder farming information, such as the General Household Survey (GHS), Income and Expenditure Survey (IES) and the Census, only capture limited dimensions of livelihoods and production information. The LCS however enables analysis on smallholder livelihoods that is otherwise not possible with other datasets separately. The sample is limited to all black households residing within the former homeland tribal areas and the sample included 634 households participating in agricultural production in some form or another. The sample was limited to the available variables enumerated and missing values systematically lowered the sample size.

*3.4 Principle Component Analysis*

The central idea behind Principle Component Analysis (PCA) involves the reduction of dimensions found in a set of data containing a large number of interrelated variables, while simultaneously retaining the maximum amount of variation in the dataset (Jolliffe, 2005). This is accomplished by transforming the data into a new dataset comprised of a new set of variables, the principle components, which are scores, calculated for the underlying dimensions in the data. These resulting components are then syntheses of the original raw data and by using these new variables will avoid the need to standardize or transform the variables for the next step in the analysis (Gaspar et al., 2008).

The analysis starts by taking p variables $X_1$, $X_2$,...,$X_p$, across *n*-households and finding combinations of these to produce a new set of indices, $Z_1$, $Z_2$,…, $Z_n$, that are uncorrelated (Manly, 1986). The first principle component is then the linear combination of the variables $X_1$, $X_2$,…., $X_p$, and is given by:

$$Z_{1i} = \alpha_{11}X_{1i} + \alpha_{12}X_{2i} + \cdots \alpha_{1i}X_{pi} \qquad (1)$$

where;

$$\alpha_{11}^2 + \alpha_{12}^2 + \cdots \alpha_{1p}^2 = 1 \qquad (2)$$

This linear combination maximizes the variance for the *X* variables amongst all such linear combinations and the coefficients are found as the eigenvectors ($\alpha$) of the sample covariance matrix (Everitt et al., 2011). The first component contributes the most to the variance as contained in the *n* number of the original variables (Essa & Nieuwoudt, 2003). The second principle component, $Z_2$, is defined as the linear combination of the original variables that accounts for the remaining variance, subject to being uncorrelated with the first principle component, i.e.;

$$Z_2 = \alpha_{21}X_{1i} + \alpha_{22}X_{2i} + \cdots \alpha_{2i}X_{pi} \qquad (3)$$

Further principle components are defined in the same way following $Z_1$ and $Z_2$. These Z-scores are uncorrelated to one another and if there are *p* variables then there can only be *p* principle components (Manly, 1986). Factors were then rotated using orthogonal (vari-max) rotation which ensures that the loadings of the variables are maximized onto one factor and minimized on the remaining factors (Field, 2009; Bidogeza et al., 2009). This rotation maximizes the sum of these variances for all of the factors (Manly, 1986). According to the Kaiser criterion, all factors exceeding an eigenvalue of one were retained and is said to be accurate if the number of variables in the analysis is less than 30 (Bidogeza et al., 2009). The eigenvalue is the sum of squared loadings for a factor and conceptually represents the amount of variance accounted for by a factor. The output from PCA is a complete new data matrix comprised of a few principle components that explain most of the dimensions in the original dataset and can now be used in the next step of the typology development.

To ease the interpretation of factors, the factor loadings output give valuable understanding of the relative contribution that an initial variable makes to a specific factor from the PCA (Field, 2009). It indicates which variables are strongly associated (correlated) with a specific factor and shows which original variables "load" onto the same factor and would identify common theme in the data (Bidogeza et al., 2009).

*3.5 Cluster Analysis*

The objective of CA in the typology development is to find optimal groupings of households which are analogous, showing a high degree of natural association within the groups and natural disassociation between groups. This is done by using the new variables from PCA which is ordered in the usual *n x p* multivariate data matrix given by the equation 4. The different values of each explanatory variable are given by **Z**, which describes each household to be clustered:

$$\mathbf{Z} = \begin{bmatrix} Z_{11} & Z_{12} & Z_{1p} \\ Z_{21} & Z_{22} & Z_{2p} \\ Z_{31} & Z_{32} & Z_{3p} \\ \vdots & \vdots & \vdots \\ Z_{ni} & Z_{ni} & Z_{np} \end{bmatrix} \qquad (4)$$

In this matrix, $Z_{ij}$ in **Z** gives the z-score of the $i_{th}$ variable on the $j_{th}$ household. The rows correspond to the variables of interest (in this case the factors form the PCA output) while the columns correspond to the different households in the data. In order to understand the closeness between different households, proximity measures are used to identify dissimilarities, similarities and distance between elements in the data (Everitt et al., 2011). In clustering households in the data, the units of proximity are usually expressed as a distance, and will be dependent on the format of the specific data used in the clustering procedure. In the case of using the components from the PCA results, the distance measure most commonly used is the Euclidean distance, which is also employed in this study. It is given by:

$$D_{12} = \left[ \sum_{k=1}^{p} (z_{1k} - z_{2k})^2 \right]^{0.5} \qquad (5)$$

Where D is the Euclidean distance between the 1st and 2nd households in the data matrix, determined by the *p* number of *z* variables, within the d-dimensional dataset. This commonly used distance function satisfies all of the conditions for a metric similarity measure according to Xu and Wunsch (2009) which include symmetry, positivity triangle inequality and reflectivity.

The Cluster Analysis conducted in this study was performed in two distinct stages. First, a hierarchical clustering method was used to create clusters of households within the sample. This method uses the similarity matrix to create a dendogram used to depict the relationships among

the different households (Anderberg, 1973). The dendogram is a two-dimensional diagram illustrating the way partitioning was done with the clustering procedure at each level and will be used to illustrate the hierarchical clustering results in this study (Everitt, 1974). The technique starts with each cluster comprising of exactly one household and combines the nearest clusters until there is only one cluster left, consisting of all of the households in the sample (Chandra & Prabuddha, 2009). This clustering method fuses individuals together which are the closest to each other and can vary in terms of the specific agglomerative techniques used. The algorithm used in this analysis was Ward's (1963) method and as mentioned earlier, the Euclidean distance measure. Ward's method encompass most of the different hierarchical clustering methods by merges chosen at each stage as to maximize an objective function which is an error sum of squares objective function.

The second, non-hierarchical method was then used which followed the abovementioned method by clustering the data units into a single classification of cluster determined by *a priori* selection (Anderberg, 1973). Using the results from the hierarchical clustering, it is possible to decide on the number of clusters in the data by referring back to the dendogram. Similar to hierarchical clustering, non-hierarchical clustering procedures have a wide range of different algorithms used. The k-means, non-hierarchical clustering method is one of the most popular and it forms clusters by specifying the number of clusters into *k* number of clusters or groups with each partition representing a cluster. Its name refers to the k-means algorithm used to calculate the mean (centroid) of each cluster (Yan, 2005). This algorithm is not based on a distance measure as used in hierarchical clustering, but uses within-cluster variation as a measure of homogeneity to segment the data so that within-cluster variation is minimized (Mooi & Sarstedt, 2011). It defines a group or prototype in terms of a centroid, usually the mean of a group of points, and it makes use of within-cluster variation measure, to create groups so that within-group variance is minimized (Kumar et al., 2006; Mooi & Sarstedt, 2011).

In order to select the number of clusters for non-hierarchical clustering, the dendogram is dissected though subjective inspection at a linkage distance where additional cluster combinations occur at a much higher distance (Kobrich et al., 2003; Mooi & Sarstedt, 2011).

However, the number of clusters must be realistic with regards to the empirical situation of the specific analysis in order for meaningful classification (Bidogeza et al., 2009).

*3.6 Cluster Validation*

According to Kobrich et al. (2003) there is no formal method to validate typologies on the basis of optimality or significance. In general, the validation of the proposed clusters can be done by using the analysis of variance (ANOVA) to test whether or not the groups differ in terms of specific quantitative variables, while the $Chi^2$ test was performed for quantitative variables in the analysis (Maseda et al., 2004; Gaspar et al., 2008; Blazy et al., 2009; Bidogeza et al., 2009; Joffre & Bosma, 2009). The ANOVA tests the null hypothesis that all the means of the specific groups are equal and the P-value would then indicate to what statistical degree one could reject the null hypothesis.

**4. Results**

*4.1 PCA results*

The Kaiser-Meyer-Olkim (KMO) (1970) measure for sampling adequacy and Bartlett's test of sphericity (Bidogeza et al., 2009) were used to test whether the dataset of 634 households and 25 variables could be used in PCA. Results from the KMO gives a value of 0.677 (>0.5), while Bartlett's sphericity test was highly significant (p-value<0.001). The selected variables in the analysis are therefore appropriate and give sufficient evidence that the selected variables can be used in PCA (Iraizoz et al., 2007). Factors with eigenvalues greater than one is retained and the 8-factors explaining 65% of the variation within the original dataset. Table 3 gives the factor loadings from the PCA results which indicate the factors that were retained and which variables are correlated to it and therefore allows for interpretation of each factor. To ease interpretation, factor loadings of greater than 0.28 is highlighted and considered to be sufficiently correlated to the specific variable.

<Table 3 here>

The first retained component from PCA, explaining 14% of the variation in the data, is a factor which correlates positively with the size of the household, the number of children in the home and income obtained from child support grants from government. Furthermore, per capita expenditure on food loads negatively on component 1 while the numbers of family farm labour utilized loads positively. This indicates that bigger households, with more children, have more available family labor for farm activities and these households spend, on average, less on food per person. Component 2 is clearly a measure which indicates the households' dependence on old age grants from government as the age, number of elderly and grant income for those above 60 years, correlate positively. This component, together with the Component 1, explains more than 27% of the variation in the data which highlights the central role that grants fulfil in determining livelihood strategies amongst farming households in the former homeland areas.

Component 3 is a measure which will give a high factor value for households with higher attachment to wage earning activities. On average, high total monthly expenditure is mostly attained through relatively higher salary incomes and therefore links closely to the number of wage earning family members residing in the household. Another variable which correlates positively with component 3 is the index for credit access with these households having access a variety of loans. This factor explains 11% of the variation in the data. Next, component 4 measures the household's education level and their asset status. This component loads positively on both the years of education of the head and the highest educated person in the household. Households with higher educational attainment are correlated with higher degrees of ownership of assets. It is clear that education plays an important role in determining the households' general welfare and ability attain and utilise more assets.

Component 5 is an index relating to the households' livestock production in terms of the number of animals (cattle, sheep, goats and chicken) the household currently have in stock. All of these variables load positively on to component 4 and correlate positively to one another

From Table 3 it is clear that component 6 correlate positively with the number of spouses living away from home and the value of household income from remittances. The negative factor loading of 0.3 on gender highlights a popular dynamic in rural South Africa where males tend to leave rural areas to find employment elsewhere in order to supply income to the family. This is therefore also a clear livelihood strategy for rural, female-headed, farming households. Component 7 loads highly on both the gender (0.64) and the marital status (0.72) of the head, both being positive. Finally, Component 8 has a high and positive factor loading for both the size of land utilized for farming and the monthly amount of inputs purchased, while it correlates negatively with the level of food insecurity. This highlights an important dynamic amongst farming households in the sense that those farming on bigger sizes of land are more food secure and able to spend more resources on farming activities.

*4.2 Results from Cluster Analysis*

The results from the Hierarchical cluster analysis are given in the dendogram in Figure 4 below. It indicates the number of cluster solutions available as one move from the top to the bottom. The cluster solution is given by dissecting the dendogram at a linkage distance of 90, at which 8 cluster groups emerge from the exercise. This step in the analysis points to relative homogenous clusters to be form in the next step, non-hierarchical clustering which will ultimately give the 8-cluster solution.

<Figure 4 here>

The non-hierarchical K-means clustering procedure yielded the results given in Table 4. This table indicates the mean values within clusters in order to understand differentiation amongst the farming household cluster. Each variable used in the analysis was then used and the ANOVA tests suggest that these groups of farming households were significantly ($p<0.01$) different from each other and therefore considered to be valid cluster groupings. Each cluster will be defined by its main characteristics and therefore explain the livelihood strategies of each household.

<Table 4 here>

Cluster 1: Younger families, married and male-headed, salary dependant

The first cluster from the analysis is typically younger families being established in rural areas. These households are characterised by having the entire family staying in the same house and income is sourced from at least 1 person working for a wage. The average salary income for these households is R868 per month, while additional income comes in the form of approximately one child grant of R263. These younger families comprise of the married, male head with his spouse and their three children, while not many elderly relatives reside in these households. Farming systems of these households are on small pieces of land of 0.51 hectares and these had the lowest animal stock numbers.

Cluster 2: Small, elderly and female headed households, dependent on an old age grant

These households are typically made up of one elderly female, which receive an old age pension from the government of R652 each month. The rest of the household consist of one adult and one child. These families had the lowest total income reported compared to the other cluster means and in association also had the lowest asset index. Furthermore, the households have the lowest amount of individuals attached to the labor market which causes these households to experience the lowest access to credit amongst all of the cluster groups. Thus, these households are typically poor, have very low access to credit, and do not own many assets. Farming activities are done on a very small scale mostly to supplement food to the household.

Cluster 3: Elderly and male-headed households' dependant on old age pensions

Households in Cluster 3 consist mainly of an elderly (70 years of age) and male (98%) household head with his elderly spouse. Amongst them they obtain more or less R1571 in old age grant income on a monthly basis. The household head has on average only 3 year of education; a typical outcome of a black South African growing up under the Apartheid system. Apart from the head and his spouse, these households often include two 2 children and a second adult living in the home. An additional income to the household comes in the form of wage income (R407) and approximately one child grant (R205). Cluster 3 households utilize bigger pieces of land (0.91 ha) compared to Cluster 1 and 2, and these also keep more animal stock of cattle (3), sheep and goats (4) and chicken (9), on average. The expenditure on farm inputs was R29 per months and 2 family members are active in farming operations.

Cluster 4: Big, female-headed households with many children, and one pensioner

The livelihood strategy of households included in this cluster is to generate income from grants. The main income source is therefore R641 coming from child grant income of approximately three children. The female household head is often the grandmother of the many children in the home and qualifies for old age pension of R 641 per month. Thus, these households are almost entirely dependent on the government social assistance programs to supply income to the household. The average household size of these households is 8 and these households often utilise more family labour (2.5) compared to the other clusters. The food security index gives the highest mean value to these households and the much lower per capita food expenditure suggests that these households are most vulnerable to hunger prevalence. This also suggest that household select into farming activities to feed the household to minimise food expenditure from more formal market channels.

Cluster 5: Remittance dependant, female-headed households

Smallholder households in Cluster 5 are made up of a relatively young female with 44 years of age and 92% of these are married. These households employ a livelihood strategy where the male spouse live elsewhere, generates income from wage earning activities which is sent back to the household in the form of a remittance transfer of R645. Staying behind in the rural areas is the female with her children which in turn receive R449 from child grants. Such livelihood decisions emanate from the problem of limited employment opportunities in rural areas and highlight the need for development to create employment.

Cluster 6: Emerging, intensive farming households producing surplus

Households in Cluster 6 are characterized by their relatively higher intensity farming system operations both in terms of labour and capital. Mostly headed by females, these households source income from both salaries (R645), a child grants (R285) and farm income as indicated in Table 5. Some additional variables indicate that 46% of the households in Cluster 6 were selling their produce which gave supplementary income of R266; the equivalent of an additional child grant in value. On average, the household size is 6 persons per household, of which 3 participate in farm labor. This was the highest number of family labor applied by the different cluster

groups. Furthermore, farming takes place on much larger land size (4.8 ha), 4 times higher than any other cluster group. Animal numbers of each category is higher than 17, while input expenditure per month was R160 per month, on average.

<Table 5 here>

These farming households therefore are willing to use household resources to invest in the production systems as a source of livelihood activity; which supplied modest levels of additional income to the household. Furthermore, Table 5 highlights another important dynamic in that these households had access to much higher levels of credit. Though lower than Cluster 7, but still significantly higher than any other cluster. This strengthens the argument that in order to scale up production, financing larger farming operations becomes essential.

Cluster 7: Employed, well-off households with married, educated heads.

Connectedness to the labor market is a key livelihood strategy for these households, with total expenditure at R11413 per month, mainly supplied by the salaries (R8944) from at least one person working for wages. Typically, these households have more educated heads with an average of 9 years of education and the highest educated individual had an average of 13 years of education. This reveals the important role that higher education play in livelihood outcomes amongst farming households. Furthermore, these households have the highest index values for asset status and the lowest probability of being food insecure. The same applies for access to credit which would be expected by their relative higher salary incomes to finance their debt. Thus, from a welfare perspective these families are significantly better-off compared to all other farming groups in the typology. Farming activities of these households tend to be part of subsistence or leisure activities on small pieces of land, and the high per capita expenditure of R368 shows that these households mostly purchase food from the formal food markets, instead of depending on their subsistence production, even though the household consisted of 5 family members.

## 5. Conclusion

This paper speaks to two shortcomings which hinder effective development of policy interventions to support smallholder farming in the former homeland areas. Developing a smallholder typology to identify dominant livelihood strategies allows for much needed analysis on diversity amongst smallholder households in rural South Africa. The analysis highlights 7 cluster groups based on the Sustainable Livelihoods classification which develops distinct livelihood strategies based on multiple factors. The typology was developed to understand these dynamics for smallholder farming households in the former homeland regions.

Results show that indeed, much diversification exist within this group of subsistence producers, each deploying specific strategies towards more resilient livelihood outcomes. Grants, both old age and child grants, affects smallholder classification which again re-emphasize the importance of these on welfare outcomes. Important to note is that even that though these received grants as main income source, they still participated in farming activities, mostly to boost household food consumption and to be less reliant on buying food from the market. Specifically Cluster 2, 3 and 4 builds their livelihoods on the receiving of grant payments from government. The latter does so by having big households with many children, while the former two clusters utilise old age pensions to support their living standard.

From a welfare perspective the results show that Cluster 2 households were the most vulnerable in terms of food security, ownership of assets and access to credit. Scaling up farm production could assist in generating sufficient food to these households. Results also indicate that those with higher education and those working for salaries are at a distinct advantage with regards to higher living standards, food consumption, higher asset attainment and lower prevalence of food insecurity. Families in Cluster 5 maintain their livelihoods by having a spouse work away from home, while the rest of the family stay in the rural areas and farms to supplement food to the household.

Cluster 6 is a group of farming households which are more inclined to farm in order to sell their produce. Such livelihoods strategies depend on the access to larger pieces of land and the ability to use income for farm inputs. This group did so by generating income from wage earning

activities and in turn managed to supplement household income with an additional R266. This group also utilized the highest number of family labor and this group invested R160 per month on farm inputs.

The results show that rural development interventions should target households listed in Cluster 6 into more efficient production systems and access to land. Such households have the highest potential to create employment, whether for family members or not. This will require sufficient access to credit and market access opportunities as these were the furthest from the market. Increasing educational attainment in rural South Africa does have major spill-over effects not only on the welfare of a select few high earning households, but also to those younger families being established in rural areas to support higher income opportunities. From a food security perspective, households such as those listed in Cluster 2 would benefit from increased agricultural support to stimulate more food availability to these households.

## 6. References

### Bibliography

Aliber, M. & Hall, R., 2012. Support for smallholder farmers in South Africa: Challenges of scale and strategy. *Development Southern Africa*, 29(4), pp.548-62.

Aliber, M. & Hart, T., 2009. Should Subsistence Farming be Supported as a Strategy to address Rural Food Security. *Agrekon*, 48(4), pp.434-58.

Anderberg, M., 1973. *Cluster analysis for applications*. New York: Academic Press.

Anderloni, L., Bacchiocchi, E. & Vandone, D., 2011. *Household financial vulnerability: an empirical analysis*. Working Paper 2011-02. Milan: University of Milan Department of Economics, Business and Statistics.

Ashley, C. & Carney, D., 1999. *Sustainable Livelihoods: Lessons from early experience*. London: Department for International Development.

Babulo, B. et al., 2008. Household Livelihood Strategies and Forest Dependence in the Highlands of Tigray, Northern Ethiopia. *Agricultural Systems*, 98, pp.147-55.

Ballas, D, Kalogeresis, T. & Labriandis, L., 2003. *A comparative study of typologies for rural areas in Europe*. Paper at the 43rd European Congress of the Regional Science Association. Jyväskylä: Centre for Computational Geography.

Barr, N., 1996. Conventional and low-input pasture improvement: A review of recent market research. *New Zealand Journal of Agricultural Research*, 39, pp.559-67.

Belsky, J., 1984. *Stratification of Hillside Farmers and some implications for Agroforestry Programs: a Case Study in Leyte, Philippines*. Masters Thesis. Faculty of the Graduate School.

Bidogeza, J., Berentsen, P., De Graaff, J. & Oude Lansink, A., 2009. A typology of farm households for the Umutara Provincein Rwanda. *Food Security*, 1, pp.321-35.

Blazy, J. et al., 2009. A Methodological Framework that accounts for Farm Diversity in the Prototyping of Crop Management Systems: Application to banana-based Systems in Guadeloupe. *Agricultural Systems*, 101, pp.30-41.

Booysen, F., 2002. Using demographic and health surveys to measure poverty: an application to South Africa. *Journal of Studies in Economics and Econometrics*, 26(3), pp.53-70.

Booysen, F.V.D.B.S., Burger, S., Maltitz, M. & Rand, G., 2008. Using an asset-index to assess trends in poverty in seven Sub-Saharan African countries. *World Development*, 36(6), pp.1113-30.

Bundy, C., 1979. *The Rise and Fall of South African Peasantry*. 28th ed. University of California Press.

Capillon, A., 1993. *Typologie des exploitations agricoles, contribution a l&etude regionale des problemes techniques*. PhD Thesis. Paris: INAP-G.

Chamberlin, J., 2008. *It's a Small World After All: Defining Smallholder Agriculture in Ghana*. Discussion Paper 00823. Development Strategy and Governance Division.

Chandra, B. & Prabuddha, S., 2009. *Food security, poverty and nutrition policy analysis: Statistical methods and applications*. Academic Press.

Chichaibelu, B., Beyene, B. & Waibel, H., 2012. *Financial vulnerability and Indebtedness of poor households in Thailand and Vietnam*. Paper presented at the Conference on International Research on Food Security, Natural Resource Management and Rural Development. Göttingen: Georg-August Universität Göttingen.

Christiaensen, L., Demery, L. & Kuhl, J., 2010. The (evolving) role of agriculture in poverty reduction: An empirical perspective. *Journal of Development Economics*, 96(2011), pp.239–54.

Cousins, B., 2010. What is a 'Smallholder'? Class Analytic Perspective on Small-scale Farming and Agrarian Reform in South Africa. *Reforming Land and Resource Use in South Africa: Inpact and Livelihoods*, 6, pp.86-111.

Cousins, B., 2013. Smallholder irrigation schemes, agrarian reform and 'accumulation from above and from below' in South Africa. *Journal of Agrarian Change*, 13(1), pp.116-39.

DAFF, 2012. *Integrated growth and development plan*. Pretoria: Directorate Communication Services National Department of Agriculture, Forestry and Fisheries.

DAFF, 2014. *Budget vote speech by the Minister of Agriculture, Forestry and Fisheries, Honourable Senzeni Zokwana*. Pretoria: Department of Agriculture, Forestry and Fisheries.

Demeke, A. & Zeller, M., 2009. *Using panel data to estimate the effect of rainfall shocks on smallholder food security and vulnerability in rural Ethiopia*. Discussion Paper No. 2/2009. Research in Development Economics and Policy.

Dorward, A., 2002. *A Typology of the Malawian Rural Households*. Working paper. Wye: Imperial College: Institutions and Economic Policies for Pro-Poor Agricultural growth.

Ellis, F., 1998. Household strategies and rural livelihood diversification. *The Journal of Development Studies*, 35(1), pp.1-38.

Emtage, N., 2004. *An investigation of the social and economic factors affecting the development of small-scale forestry by rural households in Leyte Province, Philippines: A typology of rural households in relation to small-scale forestry*. PhD Thesis. University of Queensland.

Emtage, N & Suh, J., 2005. Identifying typology of households based on the characteristics of their farming Land and livelihood systems. In *ACIAR Smallholder Forestry Project-Improving Financial Returns to Smallholder Tree Farmers in the Philippines*. pp.211-20.

Essa, J. & Nieuwoudt, L., 2003. Socio-economic dimensions of small-scale agriculture: A principal component analysis. *Development Southern Africa*, 20(1), pp.67-73.

Everitt, B., 1974. *Cluster analysis*. London: Heinemann Educational Books ltd.

Everitt, B., Landau, S., Leese, M. & Stahl, D., 2011. *Cluster analysis*. 5th ed. Londen: Wiley.

Fenyes, T. & Meyer, N., 2003. Structure and production in South African agriculture. In L. Nieuwoudt & J. Groenewald, eds. *Agriculture in the National Economy*. Pietermaritzburg: University of Natal Press. pp.21-45.

Field, A., 2009. *Discovering statistics using SPSS*. 3rd ed. Londen: SAGE Publications Ltd.

Filmer, D. & Pritchett, L., 2001. Estimating wealth effect without expenditure data - or tears: an application to educational enrollments in the states of India. *Demography*, 38, pp.115-32.

Gaspar, P. et al., 2008. Sheep farms in the Spanish rangelands (dehesas): Typologies according to livestock management and economic indicators. *Small Ruminant Research*, 74, pp.52-63.

GHS, 2013. *The General Household Survey*. Pretoria: Statistics South Africa.

Groenewald, J. & Nieuwoudt, L., 2003. Chapter 15: Demands on and Challenges for South African Agriculture. In L. Nieuwoudt & J. Kirsten, eds. *The challenge of change: Agriculture, land and the South African economy*. Pietermaritzburg: University of Natal. pp.265-83.

Hamilton, C., Mbenga, B. & Ross, R., 2012. *The cambridge history of South Africa*. 11885th ed. Cambridge: Cambridge University Press.

Iraizoz, B., Gorton, M. & Davidova, S., 2007. Segmenting farms for analysing agricultural trajectories: A case study of the Navarra region in Spain. *Agricultural Systems*, 93, pp.143-69.

Joffre, O. & Bosma, R., 2009. Typology of shrimp farming in Bac Lieu Province, Mekong Delta, using multivariate statistics. *Agriculture, Ecosystems and Environment*, 132, pp.153-59.

Jolliffe, I., 2005. *Principal component analysis*. John Wiley & Sons, Ltd.

Kaine, G. & Lees, J., 1994. *Pattern in innovation: An analysis of the adoption of practices in beef cattle breeding*. 60. Armidale : University of New England The Rural Development Centre.

Kaiser, H., 1970. A second generation little jiffy. *Psychometrika*, 35, pp.401-15.

Kirsten, J. & Van Zyl, J., 1998. Defining Small-scale farmers in the South Afrian Context. *Agrekon*, 37(4), pp.551-61.

Kobrich, C., Rehman, T. & Khan, M., 2003. Typification of Farming Systems for Constructing Representative Farm Models: Two Illustrations of the Application of Multi-variate Analyses in Chile and Pakistan. *Agricultural Systems*, 76, pp.141-57.

Kumar, P., Tan, P. & Steinbach, M., 2006. Cluster analysis: Basic concepts and algorithms. *Intoroduction to Data Miming*, pp.487-586.

Lahiff, E. & Cousins, B., 2005. Smallholder agriculture and land reform in South Africa. *IDS Bulletin*, 36(2), pp.127-31.

Laoubi, K. & Yamoa, M., 2009. A typology of irrigated farms as a tool for sustainable agricultural development in irrigation schemes. *International Journal of Social Economics*, 36(8), pp.813-31.

Larson, D., Otsuka, K., Matsumoto, T. & Kilica, T., 2014. Should African rural development strategies depend on smallholder farms? An exploration of the inverse-productivity hypothesis. *Agricultural Economics*, 45, pp.355-67.

Laurent, C. et al., 1999. Household typology for relating social diversity and technical change. The example of rural households in the Khambashe area of the Eastern Cape province of South Africa. *Agrekon*, 38, pp.190-206.

LCS, 2008/09. *Living Conditions Survey in SA*. Pretoria: Statistics South Africa.

Louw, A., 2013. Sustainable policy support for smallholder agriculture in South Africa: key issues and options for consideration. In S. Greenburg, ed. *Smallholder and agro-foodvalue chains in South Africa*. Bellville: PLAAS. pp.23-28.

Manly, B., 1986. *Multivariate statistical methods: A primier*. Londen: Chapman and Hall Ltd.

Maseda, F., Dıaz, F. & Alvarez, C., 2004. Family dairy farms in galicia (N.W. Spain): Classification by some family and farm factors relevant to quality of life. *Biosystems Engineering*, 87(4), pp.509-21.

Mooi, E. & Sarstedt, M., 2011. *A concise guide to market research: The process, data, and methods using IBM SPSS Statistics*. Springer-verlag Berlin Heidelberg.

Motsoari, C., 2012. *Access to credit and agricultural production in Lesotho*. Masters thesis. Bloemfontein: University of the Free State Department of Agricultural Economics.

Neves, D. & Du Toit, A., 2013. Rural livelihoods in South Africa: Complexity, vulnerability and differentiation. *Journal of Agrarian Change*, 13(1), pp.93-115.

NPC, 2011. *National Development Plan: Vision for 2030*. Pretoria: National Planing Commission.

Nyaga, E. & Doppler, W., 2009. Combining principal component analysis and logistic regression models to assess household level food security among smallholder cash crop producers in Kenya. *Quarterly Journal of International Agriculture* , 48(1), pp.5-23.

Oettle, N. et al., 1998. *Agricultural services in southern Africa: Encouraging sustainable smalholder agriculture in South Africa*. Glos: Environment and Development Consultancy Ltd.

Ortmann, G. & Machethe, C., 2003. Problems and Opportunities in South African Agriculture. In L. Nieuwoudt & J. Groenewald, eds. *The Challenge of Change: Agriculture, Land and the South African Economy*. Pietermaritzburg: University of Natal Press. pp.47-62.

Pauw, K., 2007. Agriculture and poverty: Farming for food of farming for money? *Agrekon*, 46(2), pp.195-218.

Perret, S., Anseeuw, W. & Mathebula, F., 2005. *Poverty and livelihoods in rural South Africa: Investigating diversity and dynamics of livelihoods*. Report on theme #1: Strategies for poverty alleviation and food security. Pretoria: University of Pretoria.

Perret, S. & Kirsten, J., 2000. *Studying the local Diversity of Rural Livelihoods Systems: An Application of Typological Techniques for intergrated Rural Development Support in the Eastern Cape (South Africa)*. Working Paper 2005-05. Pretoria: University of Pretoria, Department of Agricultural Economics, Extension and Rural Development.

Righia, E., Dogliotti, S., Stefaninic, F. & Pacinia, G., 2011. Capturing farm diversity at regional level to up-scale farm Level impact assessment of sustainable development options. *Agriculture, Ecosystems and Environment*, 142, pp.63-74.

Sahn, D. & Stifel, D., 2000. Poverty comparisons over time and across countries in Africa. *World Development*, 28(12), pp.2123-55.

Scoones, I., 2009. Livelihood perspectives and rural development. *Journal of Peasant Studies*, 36(1), pp.171-96.

Tefera, T., Perret, S. & Kirsten, J., 2004. Diversity in livelihoods and farmers' strategies in the Hararghe Highlands, Eastern Ethiopia. *International Journal of Agricultural Sustainability*, 2(2), pp.133-46.

Thompson, L., 2000. *A history of South Africa*. 3rd ed. New Haven: Yale University Press.

Tittonell, P. et al., 2010. The diversity of rural livelihoods and their influence on soil fertility in agricultural systems of East Africa – A typology of smallholder farms. *Agricultural Systems*, 103, pp.83-97.

Van Averbeke, W. & Mohamed, S., 2006. Smallholder farming styles and development policy in South Africa. *Agrekon*, 45(2), pp.136-57.

Van der Ploeg, J., 1990. *Labor, markets and agricultural production*. Westview: Boulder.

Van der Ploeg, J., 1993. Rural sociology and the new agrarian question: Perspective from the Netherlands. *Sociologia Ruralis*, 33(2), pp.240-60.

Van der Ploeg, J., 1994. Styles of farming: An introductory note on concepts and methodology. In J. van der Ploeg & A. Long, eds. *Born from within: Practice and perspectives of endogenous rural development*. Assen: Van Gorcum. pp.7-30.

Vanclay, F., Howden, P., Mesiti, L. & Glyde, S., 2006. The social and intellectual construction of farming styles: Testing dutch ideas in Australian agriculture. *Sociologia Ruralis*, 46(1), pp.61-82.

Vink, N. & Kirsten, J., 2003. Agriculture in the national economy. In L. Nieuwoudt & J. Groenewald, eds. *The challenge of change: Agriculture, land and the South African economy*. Pietermaritzburg: University of Natal Press. pp.3-19.

Visalakshi, K. & Thangavel, K., 2009. Impact of Normalization in Distributed K-Means Clustering. *International Journal of Soft Computing*, 4(4), pp.168-72.

Vyas, S. & Kumaranayake, L., 2006. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6), pp.459-68.

Ward, J., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), pp.236-44.

WFP, 2009. *Comprehensive Food Security & Vulnerability Analysis Guidelines*. Rome: World Food Program.

Wineman, 2., 2014. *Multidimentional household food security measurement in rural Zambia*. 2014 AAEA Annual Meeting, Minneapolis, MN, July 27-29, 2014. Michigan State University.

Xu, R. & Wunsch, D., 2009. *Clustering*. New Jersey: Wiley.

Yan, M., 2005. *Methods of determining the number of clusters in a data set and a new clustering criterion*. PhD Thesis. Virginia: Virginia Polytechnic Institute and State University.

## 7. Tables

*Table 1: Descriptive statistics of the variables used in principle component analysis*

| Category | Variable Name | Units | LCS (n = 634) | |
|---|---|---|---|---|
| | | | Mean | St. Dev. |
| Household characteristics | Head_age | year | 56.59 | 14.78 |
| | Head_educ | year | 5.02 | 4.37 |
| | Educ_highest in family | year | 9.10 | 3.77 |
| | Head_gender | % yes | 0.45 | 0.50 |
| | Marital Status | % yes | 0.54 | 0.50 |
| | HHsize | # | 5.25 | 2.71 |
| | HH_children | # | 2.51 | 1.90 |
| | HH_employed | # | 0.48 | 0.75 |
| | HH_elderly | # | 0.62 | 0.71 |
| | HH_spouse_away | # | 0.25 | 0.46 |
| Income and Expenditure | Income_salary | rand | 938.50 | 2433.14 |
| | Income_oldagegrant | rand | 521.69 | 657.85 |
| | Income_childgrant | rand | 281.51 | 324.01 |
| | Income_remittance | rand | 122.80 | 382.91 |
| | Income_Expenditure_total | rand | 2732.80 | 2952.36 |
| | Expenditure_percap_food | rand | 253.64 | 199.92 |
| Production Orientation | Hectares_utilized | ha | 0.86 | 2.17 |
| | Inventory_cattle | # | 2.68 | 5.63 |
| | Inventory_sheep&goats | # | 5.12 | 11.93 |
| | Inventory_chicken | # | 7.97 | 9.00 |
| | M_inputexpenditure | rand | 36.14 | 73.74 |
| | HH_farmlabour | # | 1.91 | 1.19 |
| Socio Economic | Index_foodsecurity | # | 0.14 | 0.13 |
| | Index_asset_status | # | 0.04 | 0.14 |
| | Index_credit_access | # | 0.01 | 0.05 |

Source: LCS, 2009

*Table 2: Factor loading for each of the socio-economic index variables from PCA*

| **Variable Name** | **Unit** | **Food security Index** | **Asset Index** | **Credit Index** |
|---|---|---|---|---|
| Adult hunger_always | 0=yes; 1=no | 0.024 | | |
| Adult hunger_often | 0=yes; 1=no | 0.684 | | |
| Child hunger_always | 0=yes; 1=no | 0.027 | | |
| Child hunger_often | 0=yes; 1=no | 0.684 | | |
| Food standard_less than adequate | 0=yes; 1=no | 0.249 | | |
| Distance_to_water_source | km | 0.034 | | |
| Value of dwelling | rand | | 0.288 | |
| Beds_owned | number | | 0.302 | |
| Value_vehichle | rand | | 0.223 | |
| Value_tv | rand | | 0.298 | |
| Value_refridgerator | rand | | 0.243 | |
| Value_stove | rand | | 0.363 | |
| Value_microwave | rand | | 0.258 | |
| Value_washingmachine | rand | | 0.108 | |
| Value_bed | rand | | 0.429 | |
| Value_furniture | rand | | 0.382 | |
| Value_tractor | rand | | 0.063 | |
| Value_wheelborrow | rand | | 0.143 | |
| Value_grindingmill | rand | | 0.040 | |
| Value_bicycle | rand | | 0.090 | |
| Value_cellphone | rand | | 0.229 | |
| Value_tools | rand | | 0.053 | |
| Credit_loans | installment | | | 0.005 |
| Credit_vehicle | installment | | | 0.616 |
| Credit_card | installment | | | 0.034 |
| Credit_otherbankloans | installment | | | 0.646 |
| Credit_furniture&appliances | installment | | | 0.441 |
| Credit_retail | installment | | | 0.089 |
| Credit_familymembers | installment | | | 0.026 |

Source: Own calculations

*Table 3: Factor loadings output retained from principle component analysis*

| Variable name | Principle Components | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| Head_age | -0.02 | **0.48** | 0.06 | -0.15 | 0.04 | -0.02 | -0.08 | 0.04 |
| Head_educ | -0.09 | -0.15 | -0.07 | **0.59** | -0.05 | -0.02 | 0.02 | -0.04 |
| Educ_highest in family | 0.14 | 0.03 | 0.04 | **0.58** | 0.01 | -0.08 | -0.09 | 0.09 |
| Head_gender | -0.03 | 0.03 | -0.01 | 0.00 | 0.12 | **-0.30** | **0.64** | 0.00 |
| Marital Status | 0.03 | -0.02 | 0.02 | -0.03 | -0.07 | 0.21 | **0.72** | -0.02 |
| HHsize | **0.51** | 0.07 | 0.08 | 0.03 | 0.05 | -0.02 | 0.02 | 0.02 |
| HH_children | **0.51** | 0.00 | 0.01 | -0.02 | 0.03 | 0.08 | -0.01 | -0.06 |
| HH_employed | 0.09 | -0.12 | **0.43** | -0.11 | -0.14 | -0.19 | 0.02 | 0.09 |
| HH_elderly | 0.01 | **0.57** | -0.02 | 0.03 | -0.03 | -0.04 | 0.04 | -0.01 |
| HH_spouse_away | 0.05 | -0.04 | 0.03 | -0.10 | -0.04 | **0.62** | 0.08 | 0.06 |
| Income_salary | -0.04 | -0.03 | **0.56** | 0.04 | -0.01 | -0.02 | 0.00 | -0.03 |
| Income_oldagegrant | 0.01 | **0.57** | -0.04 | 0.00 | -0.05 | 0.00 | 0.01 | 0.00 |
| Income_childgrant | **0.44** | -0.06 | -0.09 | -0.02 | -0.07 | 0.10 | -0.02 | -0.06 |
| Income_remittance | -0.02 | 0.00 | -0.05 | 0.07 | 0.03 | **0.56** | -0.04 | -0.05 |
| Income_Expenditure_total | 0.02 | 0.10 | **0.51** | 0.09 | -0.01 | 0.11 | 0.02 | -0.03 |
| Expenditure_percap_food | **-0.38** | 0.04 | 0.06 | 0.01 | 0.12 | 0.19 | -0.01 | -0.02 |
| Hectares_utilized | -0.04 | 0.03 | -0.05 | 0.07 | -0.09 | 0.02 | 0.03 | **0.70** |
| Inventory_cattle | 0.03 | -0.07 | -0.01 | -0.06 | **0.51** | -0.04 | 0.07 | 0.19 |
| Inventory_sheep&goats | -0.01 | -0.01 | -0.06 | 0.04 | **0.63** | -0.03 | -0.02 | -0.13 |
| Inventory_chicken | 0.01 | 0.05 | 0.09 | -0.01 | **0.38** | 0.18 | 0.01 | 0.05 |
| M_inputexpenditure | 0.04 | -0.05 | 0.06 | -0.13 | 0.18 | -0.02 | -0.13 | **0.55** |
| HH_farmlabour | **0.29** | 0.04 | -0.03 | 0.13 | 0.11 | 0.01 | 0.06 | 0.10 |
| Index_foodsecurity | 0.06 | -0.02 | 0.00 | -0.13 | 0.21 | -0.10 | -0.15 | **-0.30** |
| Index_asset_status | -0.07 | 0.17 | 0.16 | **0.43** | 0.14 | 0.10 | 0.03 | -0.04 |
| Index_credit_access | -0.02 | -0.05 | **0.40** | -0.11 | 0.05 | 0.02 | -0.04 | -0.05 |
| Eigenvalues | 3.54 | 3.20 | 2.63 | 1.75 | 1.48 | 1.41 | 1.20 | 1.08 |
| Cumulative explained variance (%) | 14.16 | 12.81 | 10.53 | 7 | 5.94 | 5.65 | 4.78 | 4.32 |

Source: Own calculations

*Table 4: The cluster results from k-means cluster analysis*

| | | TYPOLOGY: CLUSTERS | | | | | | | ANOVA | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Variable Name** | **Unit** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **P-value** | **F-value** |
| Head_age | years | **45.05** | **64.28** | **70.03** | 63.96 | **43.65** | **57.66** | 51.79 | 0.0000 | 92.41 |
| Head_educ | years | **7.20** | **2.78** | **2.85** | 3.00 | 5.71 | 4.59 | **9.18** | 0.0000 | 34.83 |
| Educ_highest in family | years | 9.90 | 6.42 | 9.05 | 10.01 | 8.96 | 11.21 | **13.34** | 0.0000 | 31.67 |
| Head_gender | % male | **68.63** | **23.53** | **97.67** | 20.88 | **2.38** | 44.83 | 68.42 | 0.0000 | 68.45 |
| Marital Status | % married | **62.75** | **13.07** | **96.51** | 20.88 | **92.86** | 44.83 | 81.58 | 0.0000 | 79.52 |
| HHsize | number | 4.85 | **2.93** | 5.97 | **8.12** | 5.61 | 6.52 | 5.89 | 0.0000 | 60.94 |
| HH_children | number | 2.21 | 1.00 | 2.44 | **4.66** | 3.39 | 3.07 | 2.37 | 0.0000 | 62.95 |
| HH_employed | number | **0.75** | 0.19 | 0.37 | 0.33 | 0.31 | 0.45 | **1.53** | 0.0000 | 26.29 |
| HH_elderly | number | 0.08 | **0.75** | **1.71** | 0.89 | 0.07 | 0.55 | 0.45 | 0.0000 | 142.5 |
| HH_spouse_away | number | **0.07** | 0.11 | 0.06 | 0.19 | **1.05** | 0.28 | 0.29 | 0.0000 | 99.61 |
| Income_salary | rand values | **867.89** | 131.28 | 407.01 | 250.28 | 306.30 | 645.62 | **8944.80** | 0.0000 | 249.03 |
| Income_oldagegrant | rand values | 19.80 | 652.68 | **1517.67** | 781.54 | 58.33 | 398.97 | 256.58 | 0.0000 | 133.71 |
| Income_childgrant | rand values | **263.07** | 75.16 | 205.93 | **641.98** | 449.05 | 285.52 | 121.05 | 0.0000 | 52.49 |
| Income_remittance | rand values | 35.79 | 62.58 | 47.23 | 80.24 | **545.59** | 81.49 | 85.54 | 0.0000 | 24.26 |
| Income_Expenditure_total | rand values | 2028.21 | **1398.23** | 2864.84 | 2505.64 | 2490.18 | 3140.63 | 11413.38 | 0.0000 | 147.32 |
| Expenditure_percap_food | rand values | 213.21 | 349.17 | 229.83 | **164.71** | 223.91 | 266.90 | 354.31 | 0.0000 | 13.42 |
| Hectares_utilized | hectares | 0.51 | 0.48 | 0.94 | 0.67 | 0.77 | **4.82** | **1.18** | 0.0000 | 21.13 |
| Inventory_cattle | number | 1.61 | 1.24 | 3.27 | 1.85 | 1.76 | **17.69** | **4.03** | 0.0000 | 59.07 |
| Inventory_sheep&goats | number | 2.73 | 3.54 | 4.29 | 5.04 | 4.36 | **28.90** | **6.74** | 0.0000 | 25.84 |
| Inventory_chicken | number | 5.90 | 5.70 | 8.77 | 8.53 | 8.81 | **17.83** | **13.03** | 0.0000 | 12.25 |
| M_inputexpenditure | rand values | 25.57 | 24.01 | 28.58 | 41.62 | 32.94 | **159.74** | **44.35** | 0.0000 | 17.5 |
| HH_farmlabour | number | 1.81 | **1.39** | 2.10 | 2.46 | 1.94 | **2.69** | 2.00 | 0.0000 | 11.81 |
| Index_foodsecurity | values | 0.03 | 0.05 | 0.03 | **0.08** | 0.02 | 0.02 | **0.01** | 0.0127 | 2.73 |
| Index_asset_status | values | 0.13 | 0.09 | 0.17 | 0.11 | 0.12 | **0.22** | **0.36** | 0.0000 | 30.96 |
| Index_credit_access | values | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | **0.08** | 0.0000 | 15.48 |
| Number of households | n | 153 | 153 | 86 | 91 | 84 | 29 | 38 | | |

*Table 5: The cluster results from k-means cluster analysis of additional variables*

| Variable | Unit | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 | Cluster 6 | Cluster 7 |
|---|---|---|---|---|---|---|---|---|
| Sell produce | % yes | 12.93 | 14.86 | 10.59 | 11.11 | 12.20 | 46.43 | 11.11 |
| Income from Farming | rand | 25.93 | 26.31 | 13.49 | 6.27 | 7.64 | 266.24 | 24.48 |
| Distance to Food market | km | 3.62 | 3.77 | 4.82 | 5.15 | 6.09 | 6.52 | 3.41 |
| Mean Credit repayments | rand | 71.60 | 31.78 | 73.92 | 59.14 | 77.40 | 539.10 | 1730.96 |

## 8. Figures

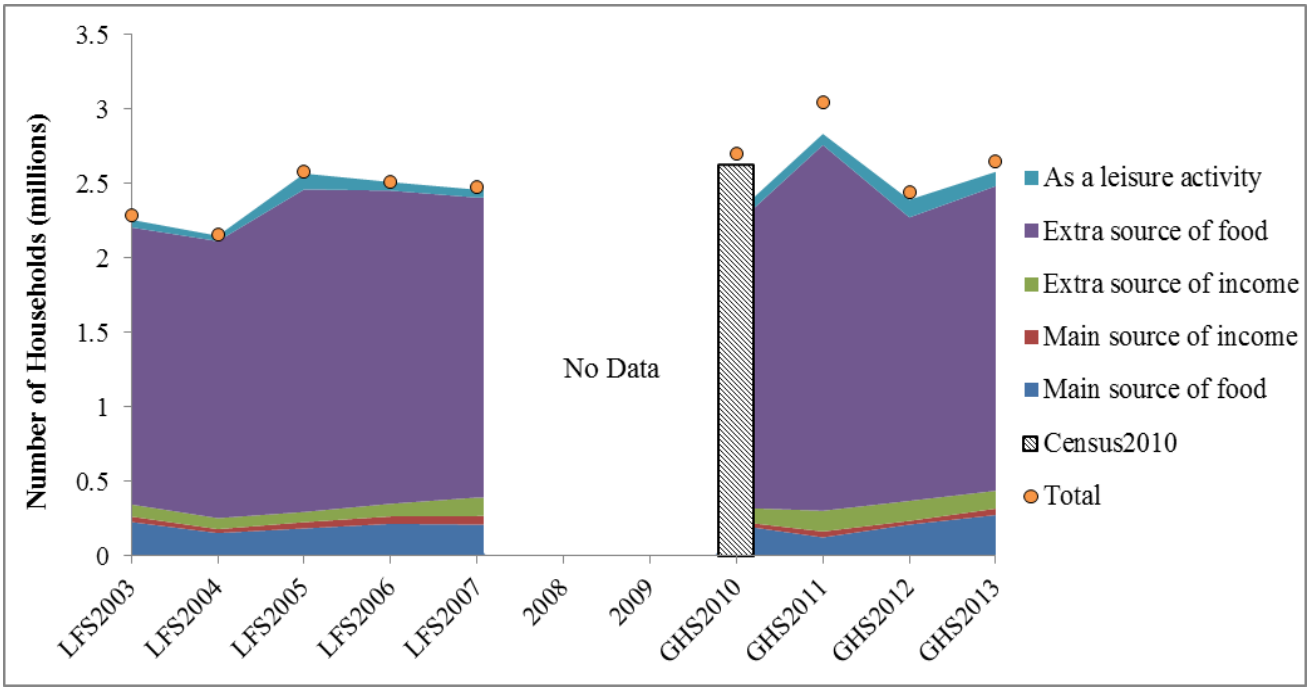


*Figure 1: Smallholder farming in South Africa*

Source: Own calculations from LFS, 2002-2007; GHS, 2009-2013; Census, 2010

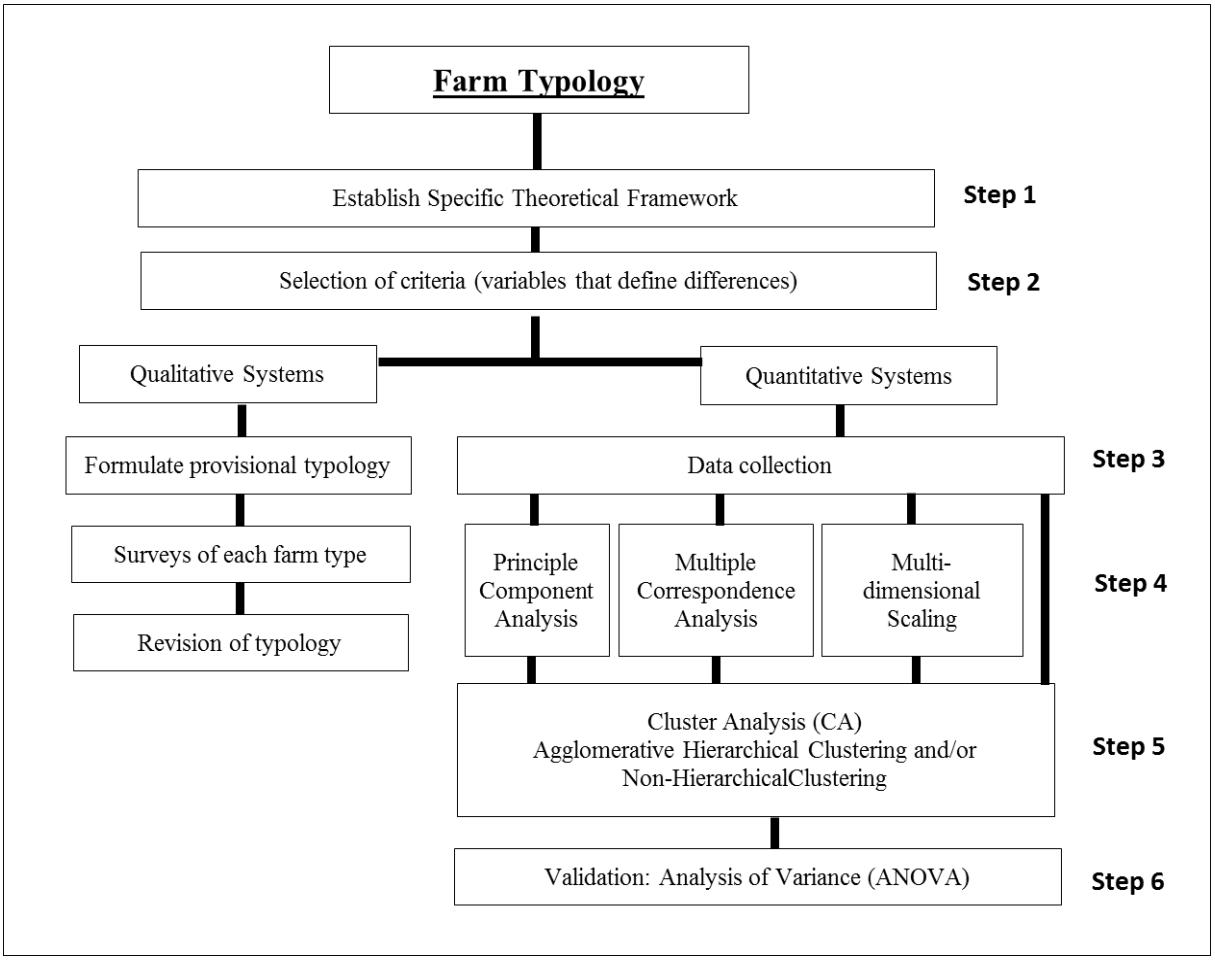


*Figure 2: Steps used to construct both quantitative and qualitative typologies.*

Source: Own compilation based on Kobric et al., 2003; Emtage et al., 2005

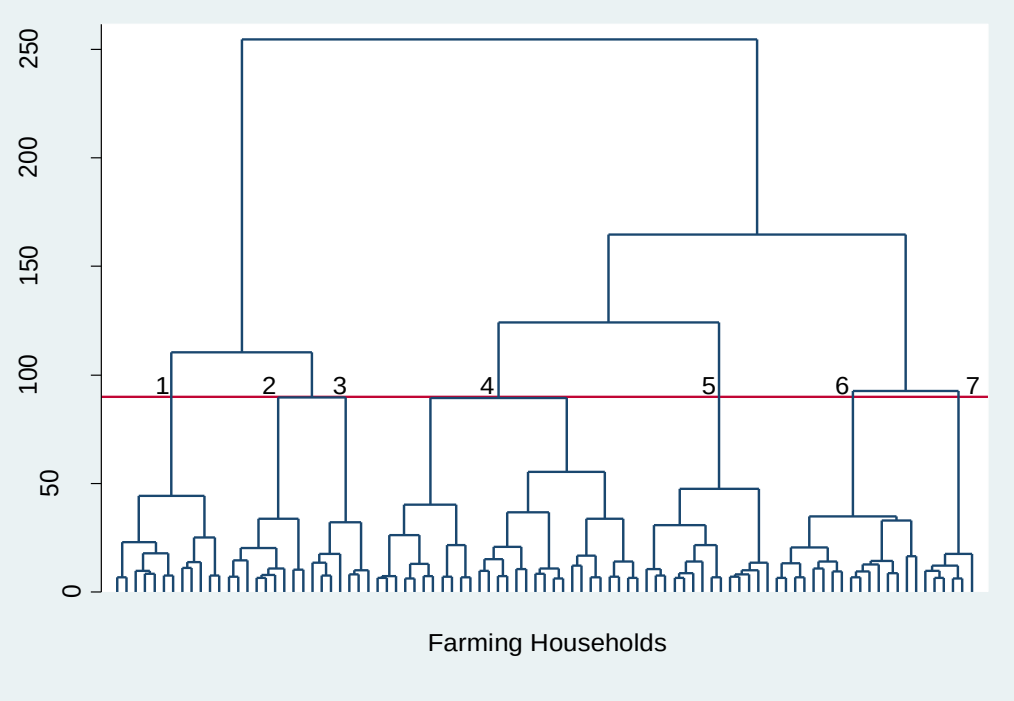


*Figure 4: Dendogram of the hierarchical clustering results*

Source: Own calculations