Types of farms for exploring diversification in Valley of Lerma (Salta-Argentina)

Setiembre de 2008

María Daniela Chavez
Instituto Nacional de Tecnología Agropecuaria (INTA EEA Salta)
Wageningen University
dchavez@correo.inta.gov.ar

Paul Berentsen
Wageningen University
Paul.Berentsen@wur.nl

Alfons Oude Lansink
Wageningen University
Alfons.OudeLansink@wur.nl

Types of farms for exploring diversification in Valley of Lerma (Salta- Argentina)

Abstract

Tobacco production has an important economic and social implication in Salta province (Argentina). There are negative sides related to tobacco production: monoculture causes environment degradation, production and price risks, and increasing restrictions to promote tobacco consumption. Diversification strategies are broadly suggested. To explore diversification, knowledge of current types of farm is needed. This paper aims to identify types of farms according to suitable variables to explore diversification in the main tobacco production area of (Valley of Lerma).

Principal component analysis and cluster analysis were applied to classify 537 farms from the data base of the National Agriculture Census (2001-2002). The selected variables for the study were determinants for diversification. Farm size, irrigation, general and specific capital goods, ownership of the land, education of the farmer, off-farm work and labor availability were used for the classification. 4 factors were selected from principal components (64.2 per cent of the total variance). These 4 factors were used for cluster analysis. Ward’s and K-means methods were applied. 6 clusters were defined. Variables showing cultivated area of crops and livestock production and an index of diversification were included for the description of the clusters. The six clusters are: Farms with scarce capital goods and less educated farmer; Small tobacco specialized farms in rent; Medium sized tobacco farms with some diversification; Diversified farms with a focus on dairy production; Large tobacco farms with a substantial level of diversification; Large, extensive calves’ production farms in property.

1 Introduction

Tobacco production entails around a quarter of the total gross value of the agricultural production of Salta province, in the Northwest of Argentina (Fittipaldi, 2004). In 2007 the production in Salta represented 33 per cent (in tons) of tobacco total production in the country. The production in the province, in terms of tons, has increased around 118% percent in the last 18 years. There are 1400 tobacco farmers and 25000 workers involved in tobacco production in Salta. About 175.000 persons depend on tobacco production for maintenance in Salta province (Cámara del tabaco, 2008). Therefore, tobacco has a relevant economic and social implication in Salta province.

There is a widely recognised negative effect of tobacco farming. Monoculture causes environmental degradation (INTA-Regional Center, 2005). Highly specialized tobacco systems involve production and price risk. Increasing restrictions to promote tobacco consumption and the matter that the government compensates the price may lead to an uncertain scenario for tobacco production. A diversification strategy is repeatedly advocated by national and provincial authorities and farmers cooperatives (Fittipaldi, 2004). Diversification means the increase of farm activities on a farm. The concept of diversification entails not only the number of farm activities but also the balance or share of them (Bravo-Ureta *et al.*, 2006; Minot *et al.*, 2006; Upton, 2004). In this paper, off-farm activities are excluded from the definition.

Farm modeling is a usable approach to explore diversification. In order to be representative a typology of tobacco farms is needed. The purpose of building typologies is to get a better comprehension of reality and to investigate technical matters in agriculture production (Kostrowicki, 1977; Landais, 1998). The variables to use in a typology depend on the aim of the research. In general, variables related to farm size, capital, labor, production pattern, soil quality and managerial ability are included to identify types of farming systems. The selected variables have to reflect the structure, functioning, objectives and restrictions of the production system (Köbrich *et al*., 2003; Paz, 1994). Quantitative techniques have been applied in current literature to build typologies to understand the variety of farming systems (Milán *et al*., 2006; Usai *et al*., 2006; Nahed *et al*., 2006).

This article aims to build a typology to identify farms for exploring diversification in Salta. The types of farms recognized in this study will be the first step in developing models concerning the analysis of diversification in the region.

This paper is organized as follows. In section 2 reasons for and determinants of diversification are explored in the literature. In section 3 the study area is described, data are presented and an explanation of the selection of variables and the methods are provided. Results are given in section 4. Discussion and conclusions are given in section 5.

2 Reasons for and determinants of diversification

2.1 Reasons for diversification

Literature shows a wide variety of reasons for diversification but all of them come down to two main reasons, namely risk reduction and improvement of income.

Risk reduction can be achieved when different sources of income have low or negative correlations. Thus, diversification of farming activities may be a way to handle risk (Minot *et al.*, 2006; Upton, 2004; Hardaker *et al*., 1997).

Income improvements can be achieved directly or indirectly. A direct improvement may arise from scope economies. The concept scope economy refers to cost savings from joint production of products compared to costs of separate production. Cost savings were found for different outputs in Germany dairy farms (Fernández-Cornejo *et al*., 1992). The shared use of inputs like labor, machinery and equipment led to costs savings in Dutch vegetable firms (Oude Lansink, 2001). Apart from scope economies, current literature reveals empirical evidence that diversification positively influences farmers' income in a direct way (Bravo-Ureta *et al*., 2006). By building scenarios Hengsdijk *et al.* (2007) found that diversification emerged as the most encouraging option to improve per capita income in traditional rice farms, compared to intensification, land expansion and exit from agriculture.

Diversification also may lead to an indirect improvement of income. For example, long distance to roads and markets can lead households to diversify into many activities to fulfill consumption needs. In this way, transaction costs are saved (Minot *et al.*, 2006; Barrett *et al.*, 2001). Another example is given by Sharma and Sharma (2005). Cost savings can be realized in continuous growing rice-wheat crop system through replacing the use of fertilizer by including a short duration legume or replacing wheat or rice by other crops, which can be considered as diversification.

Another indirect effect may arise from a shift from food production for own consumption to a cash crops production for agricultural commercialization within smallholders (Minot *et al.*, 2006).

2.2 Determinants for diversification

Determinants define the suitability of a farm for diversification. Land area, irrigation, capital goods, age, education level, off-farm work and labor availability can be determinants for diversification.

Total area of land is important in the case of arable farms. There is empirical evidence from current literature that land has a positive effect on diversification (Bravo-Ureta *et al*., 2006; Benin *et al.*, 2004). Larger area of land may motivate a farmer to devote part of it to introduce diversification in the farm.

Irrigation may have an influence on the decision to diversify in the farm. Empirical analysis showed a positive relation between irrigation and tobacco cultivated area at household level in India. The results suggest that irrigation does not encourage farmers to diversify (Panchamukhi, 2000).

The type of capital goods may have opposite effects on diversification. Specific capital goods may contribute to output specialization whereas general capital goods may facilitate diversification. For example general machinery can be used more efficiently among activities placed at different time of the year (Fernández-Cornejo *et al*., 1992; Hardaker et al. (1997). It can be expected that the availability of specific capital goods like tobacco stoves, backpacks, grain machinery, pasture machinery will prevent farmers to shift to diversification. Conversely, general capital goods like tilling tools, tractors, sprayers and fertilizer machinery, trucks and barns can motivate farmers to diversify.

Empirical data reveal highly positive effect of land tenure on output diversification in Central America, suggesting that owners grow a higher number of production items (Bravo-Ureta *et al*., 2006). Someone that relies on rented land to produce will be limited in the decisions regarding land management (Caballero, 2001). The owner of the land may be more disposed to experiment new activities to improve income in a medium or long run. Conversely, a farmer that rents the land may focus on getting profits in the short run.

The age of the farmer may affect diversification decisions. Empirical research found that the number of crops increase with the age of farmers in Vietnam, suggesting that they try new crops as they increase their experience along their lives (Minot *et al*., 2006). The same was found within more diversified farms in West Midlands (United Kingdom). Farmers involved in more diversified farms have significant farming experience; a survey showed that 70 percent of them were older than 45 years (Ilbery, 1991).

Education level has a strong and positive influence on the number of grown crops, stressing the importance of education and ability to understand information coming from extension services or other sources (Minot *et al.*, 2006). Bravo-Ureta (2006) found a positive effect of the average level of education for household members on diversification in Central America.

Off- farm work may influence the decision to diversify. A farmer who works also outside the farm probably will be less disposed to be involved in many different production activities due to a lack of time. Results of an empirical study suggest that farmers more occupied in other activity than agriculture are less expected to diversify in high value crops because of lack of time and skills (Birthal *et al.*, 2006).

Labor will be used more efficiently if it can be allocated along the year in a mix of activities (Hardaker *et al*., 1997). Economies of scope can arise from sharing labor for different outputs. Empirical data suggest that diversification in high-value crops is concentrated among households having enough labor supply (Birthal *et al.*2006).

3 Data and method

3.1 Study area

The Valley of Lerma is an elongated plain between mountains which is located in the province of Salta, Northwest of Argentina, between parallels of 24º 30'and 25º 38' (Southern latitude) and between meridians 65º 22'and 65º 37' (Western longitude). It is 120 km long (North-South orientation) and 25 km wide at the central part. It comprises an area of approximately 2400 km^2 (Baudino, 1996). It has a temperate climate and the annual rainfall varies from 500 to 1000 mm. Rain fed agriculture is feasible on the most humid parts of the valley. Irrigation is needed to compensate water shortage in winter and spring times in the less humid parts. Tobacco is grown on irrigated land (Bravo *et al*., 1999). Next to tobacco as the main crop vegetables, bean, corn, fruits, pastures, beef and milk cattle are produced in the area.

The Valley of Lerma consists of parts of 7 geographical departments which are called Capital, Cerrillos, Chicoana, Guachipas, La Caldera, La Viña and Rosario de Lerma (INTA, 2005). Tobacco cultivated area is mainly concentrated in the central part of the valley. Cerrillos, Chicoana and Rosario de Lerma departments produce 73 per cent of the total production in tons of Salta (Cámara del tabaco, 2008). For this reason, data from these three departments are used to build a typology.

3.2 Description of data

The source of data for this study was the Agricultural Census carried out by the National Institute of Statistics and Census (INDEC) in 2002. The reference period of the census comprises July 1st, 2001 to June 30th, 2002. The reference date for stock variables like number of heads and facilities and machinery is June 30th, 2002. To summarize, the variables show general information about the farm and the farmer, use of land, technical practices, stock of different cattle, inventory of buildings, facilities, machinery, equipment and vehicle, permanent and temporary labor, way of management and marketing channels.

Data were recorded following the departments borders. Since the natural borders of the Valley of Lerma do not fit the borders of the departments, it was necessary for this analysis to select the farms that are included in the study area. Experts identified those variables that helped to include farms in the valley. These variables were irrigated land, tobacco cultivated land, tobacco stoves, dairy cattle, beef cattle, legumes cultivated land, alfalfa cultivated land, Buffel grass cultivated land, vegetables cultivated land, maize cultivated land and wheat cultivated land. This selection resulted in 641 observations. These observations were then checked on important missing values which led to a final usable number of observations of 537.

3.3 Selected variables

The selected variables are developed from the original variables in the data base. The selected variables concern the determinants for diversification (section 2). In total, 16 variables are included to identify types of farms for exploring diversification. The variables description and statistics are presented in Table 1.

Table 1 Selected variables to be used for principal components

Name of the variable [a)]	Description	Mean	St.deviation	Maximum
Farm size				
Total land	Hectares	126.08	326.96	4600
Irrigation				
Irrigated area	Percentage of total land	47.97	36.89	100
General capital goods				
Tractors	Number	1.92	2.39	16
Tilling tools	Number	2.86	2.41	17
Trucks and other vehicles	Number	2.21	3.63	37
Fertilizer machinery	Number	0.29	0.60	6
Sprayers	Number	0.50	1.09	14
Barns	Number	1.72	2.01	12
Specific capital goods				
Tobacco stoves	Number	5.76	9.87	86
Backpacks for spraying	Number	2.75	3.87	30
Grains machinery	Number	0.34	0.71	5
Pastures machinery	Number	0.22	0.82	6
Ownership of land				
Land in property	Percentage of total land	61.98	45.48	100
Education				
Education level of the farmer	=1 more educated =0 less educated	0.47	0.5	1
Off-farm work				
Works outside the farm	=1 works =0 does not work	0.19	0.40	1
Labor availability				
Permanent workers	Number	3.31	5.57	52

a) minimum value for all variables= 0

Total land includes not only the cultivated land but also natural forests and pastures land, suitable (but not used) land, not suitable land and land devoted to houses, roads, barns, etc. The binary variable level of education of the farmer/s takes the value 1 in case that the farmer /s or other person in the farm has at least completed secondary school. It takes the value of 0 in case of less study than complete secondary school. The binary variable farmer/s works outside the farm takes the value of 1 when the farmer works outside and 0 when the farmer works in the farm exclusively. The variable permanent workers includes the number of workers that work every day during six or more months per year in the farm. Age was not included due to a high percentage of missing data.

3.4 Principal components analysis

The objective of principal components analysis is the reduction of the dimensionality of multivariate data. The data have to be highly correlated to give reason for applying principal components. The suitability of the original data for dimension reduction has to be tested. The sphericity test developed by Bartlett tests the null hypothesis that the correlation matrix of the population is the identity matrix (a perfectly spherical set of data) and then, data are independent. If the null hypothesis can be rejected it may be justified to use principal components for data reduction. Kaiser-Meyer-Olkin (KMO) measure of Sampling Adequacy indicates the amount of variance in the variables that might be caused by principal factors. High values, close to 1, suggest that a factor analysis may be useful and values less than 0.5 indicates the analysis is not helpful (Lattin *et al.*, 2003; SPSS, 2005).

In principal components, the original variables are linearly combined in new variables which are called components. The first components explain as much of the available information as possible. Each component is uncorrelated with each other. There are different criteria that can be followed to decide the number of components to be retained. In this research, Kaiser´s rule is followed. This criterion suggests keeping principal components with eigenvalues (variance of each component) larger than one (Lattin *et al.*, 2003; Köbrich *et al.*, 2003). The retained components are used in cluster analysis to determine types of farms to explore diversification.

3.5 Cluster analysis

Cluster analysis entails the division of a large group of observations into smaller and more homogeneous groups. Hierarchical and partitioning techniques for clustering are applied in this study. Hierarchical methods result in a tree structure where the *k*-cluster solution is built by the union of two clusters from the k + 1 cluster solution. In partitioning methods observations are separated into a number of subgroups and the *k*-cluster solution and the k + 1 cluster solution are not necessarily combined. Both methods provide a solution where each observation is allocated to only one cluster and all objects are assigned to some cluster (Lattin *et al.*, 2003). The reason to include both techniques is that they are complementary clustering methods. While the hierarchical method does not require a previous knowledge of the number of clusters, the partitioning needs to establish a prior number of clusters. The hierarchical method is applied in an exploratory way and the solution is used in a partitioning method to refine the cluster solution (Sharma, 1996).

First, the hierarchical method in an agglomerative way is applied in this study. The agglomerative approach starts with each observation in a single cluster and at following steps clusters are joined, until only one cluster contains all the observations. The graphical result of these steps is called dendrogram, which is a hierarchical tree structure (Lattin *et al.*, 2003; Köbrich *et al*, 2003). The agglomeration schedule is another result of the hierarchical method. It shows the clusters combined at each stage and the distances between them. A sudden jump in this distance suggests that a good solution was found. A good number of clusters will be the one before the gap (SPSS, 2005). There are different methods for clustering within the hierarchical agglomerative approach. In this research Ward's method is applied. Ward's method seeks to achieve clusters with the smallest within cluster sum of squares. Agglomerative clustering does not give a definitive number of clusters to represent the data. Different numbers of clusters can be found at different distances in a dendrogram and the final decision involves a substantial quantity of subjectivity (Lattin *et al.*, 2003). A partitioning method which is the K-means clustering is applied following the hierarchical method. The goal of K- means method is to split the total number of observations into a predetermined number K of groups. K-means clustering requires an initial partition of the data into K clusters. In this research the number of clusters came from the previous step (Lattin *et al.*, 2003; Valeeva *et al.*, 2005). Kruskal-Wallis test to prove the differences between the clusters is provided.

4 Results

4.1 Principal components

KMO test and Bartlett's test were performed to test the suitability of the data for a dimension reduction. KMO test equals to 0.896 and a high significance of Bartlett's test (p-value equals to 0.000) to reject the hypothesis of sphericity of multivariate data, suggest the data are suitable for applying principal component analysis.

Principal components analysis was applied on the 16 selected variables (Table 1). Following Kaiser's rule, 4 components were selected. Table 2 shows the variance explained by the 4 extracted components.

Table 2. Principal components analysis. Total variance explained by 4 components

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	6.268	39.177	39.177
2	1.864	11.647	50.825
3	1.128	7.051	57.876
4	1.011	6.322	64.198

Total column shows the amount of variance in the original variables accounted for by each component (eigenvalue). The column of percentage of variance presents the ratio of the variance accounted for by each component to the total variance of the entire variables. The cumulative column explains the percentage of variance accounted for by n components. The 4 components explain 64.198 % of the variance in the original variables. These components can be used to reduce the complexity of the data with a 35.80 % loss of information.

Table 3 presents the correlations (loadings) between the extracted 4 components and the original variables.

Table 3 Rotated component matrix

Variables	Component 1	2	3	4
Total land	.215	-.046	.569	.368
Irrigated land	.213	.056	-.838	-.038
Tractors	.883	.329	.025	-.056
Tilling tools	.603	.480	.171	-.207
Trucks and other vehicles	.819	.007	.027	.083
Fertilizer machinery	.728	.219	-.059	.091
Sprayers	.507	.473	.009	.053
Barns	.756	.183	-.074	-.033
Backpacks for sparring	.724	.068	-.185	-.072
Tobacco Stoves	.895	.012	-.041	-.030
Grain machinery	.475	.579	.196	-.113
Pastures machinery	.059	.859	.005	.125
Land in property	-.025	.186	.748	-.110
Education level	.247	.360	-.090	.452
Work outside	-.193	-.023	.098	.788
Permanent workers	.758	.293	.098	.033

Extraction method: Principal Component Analysis
Rotation method: Varimax with Kaiser Normalization

The first component explains 39.177 per cent of the variance. It is positively and highly correlated with general capital goods and specific capital goods related to tobacco production like stoves and backpacks, and permanent workers. This factor stresses the importance of tobacco production. The second component (11.647 per cent of variance) is related to specific capital goods like grain machinery and pastures machinery and it shows the importance of other types of production. The third component (7.051 per cent of variance) is correlated with total land, land in property and irrigated land (negative in this case) and it represents the size and ownership of the farm. The last component only explains 6.322 per cent of the variance and it has correlation with education level and the work outside the farm, suggesting the importance of outside work.

4.2 Cluster analysis

The four components were used for cluster analysis. First, Ward’s method was applied. A gap of the distance coefficient in the agglomeration schedule indicates a possible solution of 6 clusters. After analyzing the dendrogram (not shown here because of its huge length) and the agglomeration schedule and considering the high numbers of observations included in the study, 6 clusters were selected. Thus, 6 clusters were used in the application of the K-means method. The size of the clusters obtained by the two techniques (Ward´s and K-means) can be compared. Correspondences and differences are found. There are two larger clusters of more than 100 farms in each case. Ward´s method presents two clusters of 177 and 122 farms respectively while K-means shows one of 195 and other of 205. Ward´s method gets two clusters of 93 and 94 farms while K-means gets one of 90 farms. The smallest clusters have 43 and 8 farms in the case of Ward´s method while K-means presents a cluster of 38 and other two of 6 and 3 members.

Table 4 presents the farm types following from the 6 clusters (K-means method). Next to the initially selected variables other variables are presented, like cultivated area of different cash crops and pastures, number of heads of different livestock and a diversity index to assess current level of diversification of the clusters. The diversity index is a scalar built from the area shares allocated to crops (including those crops devoted to livestock production) and it shows both number of crops and their relative abundance (Benin *et al*., 2004). The so-called Simpson Index of diversity ranges between 0 and 1. The value of the index is 0 in case of complete specialisation and has values closer to 1 as the number of crops increases. The index is calculated as follows:

$$SID = 1 - \sum_{i=1}^{n} P_i^2$$

Where SID is the Simpson Index of Diversity and P_i is the proportionate area of i^{th} crop in the total cropped land (Joshi *et al*., 2003). Crops include cereals, tobacco, crops for seed production, legumes, annual pastures, perennial pastures, vegetables, flowers, aromatics, fruits, cultivated forests, nurseries, natural pastures and natural forests.

Table 4. Comparison of selected clusters. K-means method

Variables [a)]	*Cluster 1* *N= 205*	*Cluster 2* *N= 195*	*Cluster 3* *N=90*	*Cluster 4* *N=38*	*Cluster 5* *N= 6*	*Cluster 6* N=3
Farm size						
Total land (ha)	147.12	21.95	134.36	223.12	510.33	3211.33
Irrigation						
Irrigated area (% of land)	19.54	73.30	56.45	55.42	45.27	0.35
General capital goods						
Tractors (n°)	0.76	1.16	4.54	4.34	12.33	0.67
Tilling tools (n°)	2.30	1.81	4.82	6.00	7.50	1.33
Trucks and other vehicles (n°)	0.87	1.46	4.63	3.82	26.50	1.67
Fertilizer machinery (n°)	0.02	0.14	0.89	0.79	2.00	0.00
Sprayers (n°)	0.12	0.22	1.27	1.87	2.17	0.00
Barns (n°)	0.84	1.29	3.71	2.92	8.50	1.33
Specific capital goods						
Tobacco stoves (n°)	1.42	3.78	15.58	8.00	59.33	0.00
Backpacks for spraying (n°)	1.07	2.26	6.37	4.13	14.50	0.00
Grains machinery (n°)	0.14	0.06	0.93	1.37	1.67	0.00
Pastures machinery (n°)	0.01	0.02	0.09	2.79	0.00	0.00
Ownership of land						
Land in property (%)	93.33	20.51	69.26	84.79	66.12	100
Education						
Education level of the farmer	0.27	0.50	0.64	0.92	0.67	1.00
Off-farm work						
Works outside the farm	0.13	0.31	0.00	0.32	0.17	1.00
Labor availability						
Permanent workers (n°)	1.46	1.66	7.50	8.16	27.67	1.33
Current level of diversification						
Index of diversification (n°)	0.28	0.13	0.33	0.62	0.53	0.02
Crop production						
Cereals (ha)	0.71	0.46	6.15	4.47	8.33	0.00
Tobacco (ha)	2.66	10.10	43.29	20.80	169.17	0.00
Legumes (ha)	3.06	1.32	30.98	19.00	155.50	0.00
Anual pastures (ha)	3.83	0.63	3.41	36.18	12.50	13.00
Perennial pastures (ha)	1.67	1.06	4.80	41.07	13.33	5.33
Vegetables (ha)	0.95	1.01	1.30	2.20	11	0.00
Livestock production						
Calves (n°)	6.56	0.03	1.24	25.66	39.33	134.00
Fatten livestock (n°)	5.11	0.00	2.19	39.68	40.33	5.00
Dairy livestock (n°)	3.13	1.52	4.04	100.55	0.00	0.00

Kruskal-Wallis non parametric test: all variables significant at 0.01 level

Cluster 1. Farms with scarce capital goods and less educated farmer

This cluster is the largest and it represents 38.2 per cent of the total farms. Approximately 93 per cent of the land is in property. The average total land is high; however the cultivated land is very low. This contrast suggests that farms have an important percentage of not suitable land, or natural forests and pastures or that they do not have enough capital to produce. Only 20 per cent of the land has irrigation. They present some level of diversification (index of diversification equals to 0.28). They produce less than 3 hectares of tobacco on average. General and specific capital goods are limited in terms of their number. The farmer has the lowest level of education of all and in general does not work outside the farm.

Cluster 2. Small tobacco specialized farms in rent

This cluster represents 36.31 per cent of the farms. The total land is the lowest of all. In terms of the irrigated land the percentage is the highest (73.30%). On average 21 per cent of the land is in property, what means that most of the farmers rent the land to produce. They are highly specialized in tobacco, still they are small tobacco farms. The index of diversification is low (indicating specialization) and the area devoted to other crops is low. The level of general capital goods is not high. The farmer has a medium education level and he/she can have an off work in some cases.

Cluster 3. Medium sized tobacco farms with some diversification

This group comprises 16.76 per cent of the farms. The total land is on average the second smallest of the six clusters. 56 per cent of the land has irrigation. The index of diversification is 0.33 suggesting some level of diversification. They are medium sized tobacco farms (average of 43 ha). Next to tobacco these farms also produce legumes and (in lower proportion) cereals, pastures, vegetables. They have livestock production but not important. The farmer is in general educated and he/she works exclusively at the farm.

Cluster 4. Diversified farms with a focus on dairy production

This group represents 7 per cent of the farms. Land is mainly in property. 55 per cent of the land has irrigation. The number of machinery to produce pastures is the highest of the six clusters. The index of diversification takes the highest value of all the clusters (0.62). Farms in this type have the largest pastures (annual and perennial) cultivated land of all and the highest number of dairy livestock. Also fatten livestock is important. The education level of the farmer is high and he/she can work outside the farm.

Cluster 5. Large tobacco farms with a substantial level of diversification

This group accounts for only 1.1 per cent of the total farms. The size of the farm is the second of all (on average 510.33 ha). Around 45 per cent of the land has irrigation. Except for pasture machinery they have the highest number of general and specific capital goods. Permanent workers are significant higher than the rest of the groups. The index of diversification is 0.53 and it reveals an important level of diversification compared with other groups. This group is the most important tobacco producer. Legumes are the second most important crop in terms of grown area. They also produce cereals, pastures, vegetables, calves and fatten livestock. The farmer has good level of education and in general works only in the farm.

Cluster 6. Large, extensive calves' production farms in property

This group represents only 0.56 per cent of the total. Farms are the largest of all. Land is totally in property, with low percentage of irrigated land. The number of capital goods is low. The index of diversification is the lowest of all, suggesting a high level of specialization. The education level is high and all the farmers work also out of the farm. They are specialized in livestock production, as it is suggested by the highest number of calves of all and the absence of cash crops. The education level of the farmer is the highest of all and he/she also works outside

5 Discussion and conclusion

Tobacco production has an important economic and social implication in Salta province. There are negative sides related to tobacco production: monoculture causes environment degradation, production and price risks, and increasing restrictions to promote tobacco consumption. Diversification strategies are broadly suggested. To explore diversification knowledge of current types of farm is needed. This paper aimed to identify types of farms according to suitable variables to explore diversification in Salta.

Principal component analysis and cluster analysis were applied to classify 537 farms The selected variables for the study were determinants for diversification. Farm size, irrigation, general and specific capital goods, ownership of the land, education of the farmer, off-farm work and labor availability were used for the classification. 4 factors were selected from principal components accounting for 64.2 per cent of the total variance. These 4 factors were used for cluster analysis. Ward's and K-means methods were used. 6 clusters were selected.

The first cluster is the largest (it accounts for 38.2 per cent of total observations). The farmer is in general the owner of the land and devotes most of the time to the farm suggesting that he/she does not have another income source. He/She has the lowest level of education of all the types. One point that may prevent to improve diversification is the low availability of capital goods. This group probably needs to improve its income and in this way, diversification is interesting to be explored. Yet, capital goods (and likely cash capital) can be an important constraint. The second type accounts for the 36 per cent of all. Farms are highly specialized in tobacco. The low (on average) percentage of land in property suggests that this type includes renters of the land. The fact that they are not owners of the land may have policies implications. Medium to long term tenancy contracts may consider the obligation of including rotations to improve soil quality and to prevent soil degradation. The third cluster accounts for around 17 per cent of the total. This group accounts for medium tobacco farms with the presence of legumes and other crops and pastures. An improvement of diversification to improve soil quality may be also interesting here. The fourth cluster accounts for the 7 per cent of the total. This group shows the highest level of diversification (average of 0.62). In this group livestock production is evident. Farmers grow pastures of quality to feed dairy and fatten livestock. As tobacco farmers they have a small to a medium size of production. The presence of perennial pastures (like alfalfa) suggests some kind of good soil management. The fifth cluster is the second smallest of all (only 1.1 per cent). The level of diversification is important in comparison with others. This group of farmers includes the largest tobacco and legumes producers. It is likely that diversification to improve income is not applicable in this case. But diversification in the sense of including rotations to improve soil quality can be explored. The last cluster is the smallest (only 0.56 per cent of observations). Land availability is the highest but it seems to be with natural forests and pastures. They grow a few hectares of pastures of low quality to feed calves at an early age. The low availability of irrigated land and capital goods can be a barrier to improve diversification.

The second cluster is one of the most representatives in terms of the number of farms. This group shows interesting possibilities to explore diversification. The issue that the farmer is renting the land suggests the possibility to include rotations to maintain soil quality. Cluster 3 is the second in terms of the hectares of tobacco production of all. Diversification to improve soil quality can be explored.

References

Barrett, C.B., Reardon, T. and Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications. *Food Policy* 26: 315-331

Baudino, G. (1996). Hidrogeología del Valle de Lerma, provincia de Salta. Doctoral Thesis. National University of Sata, 166 p.

Benin, S., Smale, M., Pender, J., Gebremedhin, B. and Ehui, S. (2004). The economic determinants of cereal crop diversity on farms in the Ethiopian highlands. *Agricultural Economics* 31: 197-208

Birthal, P.S., Jha, A.K, Joshi, P.K, and Singh, D.K. (2006). Agricultural diversification in North eastern region of India: Implications for growth and equity. *Indian Journal of Agricultural Economics* 61 (3): 328-340

Bravo, G., Bianchi, A., Volante, J., Alderete Salas, S., Sempronii, G., Vicini, L., Fernández, M., Lipshitz, H. and Píccolo, A. (1999). Regiones Agroeconómicas del Noroeste Argentino. Primeras Jornadas de SIG. INTA-UNSE-Gobierno Santiago del Estero. CD.

Bravo-Ureta, B.E,, Solís, D., Cocchi, H. and Quiroga, R.E. (2006). The impact of soil conservation and output diversification on farm income in Central American hillside farming. *Agricultural Economics* 35: 267-276

Caballero, R. (2001). Typology of cereal-sheep farming systems in Castile-La Mancha (south-central Spain). *Agricultural Systems* 68: 215-232

Cámara del Tabaco de Salta (2008). La producción de tabaco en la República Argentina, 16 p.

Fernández-Cornejo, J., Gempesaw, C.M., Elterich, J.G. and Stefanou, S.E. (1992). Dynamic Measures of Scope and Scale Economies; An Application to German Agriculture. *American Journal of Agricultural Economics* 74: 329-242

Fittipaldi, C.R. (2004). Análisis de una estrategia de diversificación de la empresa tabacalera. Tesis de maestría. Escuela de Negocios. Universidad Católica de Salta, 153 p.

Hardaker, J.B., Huirne, R.B.M and Anderson, J.R. (1997). Coping with risk in agriculture. Cab International.

Hengsdijk, H. , Guanghuo, W., Van den Berg, M.M., Jiangdi, W., Wolf, J., Changhe, L., Roetter, R.P. and Van Keulen, H. (2007). Poverty and biodiversity trade-offs in rural development: A case study for Pujiang county, China. *Agricultural Systems* 94: 851-861

Ilbery, B.W. (1991). Farm diversification as an Adjustment Strategy on the Urban Fringe of the West Midlands. *Journal of Rural Studies* Vol 7 (3): 207-218

INDEC (2002). Instituto Nacional de Estadística y censos. Censo Nacional Agropecuario. Data base

INTA- Regional Center (2005). Instituto Nacional de Tecnología Agropecuaria. Salta-Jujuy. Plan Tecnológico Regional.

INTA (2005). Instituto Nacional de Tecnología Agropecuaria. Laboratorio de teledetección. http://www.inta.gov.ar/prorenoa/zonadescarga/regiones_agroeco/mapa_agroeco_salta_jujuy_1.pdf. February 22, 2007

Köbrich, C., Rehman, T. and Khan, M. (2003). Typification of farming systems for constructing representative farm models: two ilustrations of the application of multi-variate analices in Chile and Pakistan. *Agricultural Systems* 76: 141-157

Kostrowicki, J. (1977). Agricultural typology concept and method. *Agricultural Systems* 2: 33-45

Landais, E. (1998). Modelling Farm Diversity. New approaches to Typology Building in France. *Agricultural Systems* 58 (4): 505-527

Lattin, J.M., Carroll, J.D. and Green, P.E (2003). Analyzing multivariate data. Duxbury applied series. Thomson. Brooks/Cole, 525 p.

Milán, M.J., Bartolomé, J., Quintanilla, R., García-Cachán, M.D., Espejo, M., Herráiz, P.L., Sánchez-Recio, J. M. and Piedrafita, J. (2006). Structural characterisation and typology of beef cattle farms of Spanish wooded rangelands (dehesas). *Livestock Science* 99: 197-209

Minot, N., Epprecht, M., Anh, T.T.T and Trung, L.Q.T (2006). Income diversification and poverty in the Northern Uplands of Vietnam. Washington, D.C. International Food Policy Research Institute. Research Report Nº 145, 149 p.

Nahed, J., Castel, J.M, Mena, Y. and Caravaca, F. (2006). Appraisal of the sustainability of dairy goat systems in Southern Spain according to their degree of intensification. *Livestock Science* 101: 10-23

Oude Lansink, A. (2001). Long and Short Term Economies of Scope in Dutch Vegetable Production. *Journal of Agricutural Economics* Vol 52 (1): 123-138

Panchamukhi, P.R (2000). Agricultural diversification as a tool of tobacco control. The World Health Organization (WHO) International Conference on Global Tobacco Control Law: Towards a WHO Framework Convention on Tobacco control, 48 p.

Paz, R. (1994). Estructura agraria y sistemas de producción en el noroeste argentino. Su determinación a partir del análisis multivariante. *Población y Sociedad* 2: 81-102

Sharma, S. (1996). Applied Multivariate Techniques. John Wiley & Sons, Inc., 493 p.

Sharma, S.K. and Sharma, S.N. (2005). Effect of Crop Diversification of Rice-Wheat Cropping System on Productivity and Profitability. *Journal of Sustainable Agriculture* Vol 26 (1): 39-48

SPSS (2005). SPSS 14.0.1. Tutorial: Factor analysis and hierarchical cluster analysis. Chicago, US.

Upton, M. (2004). The role of livestock in economic development and poverty reduction. Pro-poor Livestock Policy Initiative (PPLPI). Food and Agriculture Organization (FAO). Working paper Nº 10, 57 p.

Usai, M.G., Sara Casu, Molle, G., Decandia, M., Ligios, S. and Carta, A. (2006). Using cluster analysis to characterize the goat farming system in Sardinia. *Livestock Science* 104: 63-76

Valeeva, N., Meuwissen,.M, Bergevoet, R., Oude Lansink, A. and Huirne, R. (2005). Improving food safety at the dairy farm level: farmers' and experts' perceptions. *Review of Agricultural Economics* 4: 574-592