ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 07/2004

The Estimation of Different Technologies using a Latent Class Model

Julio del Corral y Antonio Álvarez

Departamento de Economía

Universidad de Oviedo

Available online at: www.uniovi.es/economia/edp.htm

UNIVERSIDAD DE OVIEDO

DEPARTAMENTO DE ECONOMÍA

PERMANENT SEMINAR ON EFFICIENCY AND PRODUCTIVITY

ESTIMATION OF DIFFERENT TECHNOLOGIES USING A LATENT CLASS MODEL[*]

Julio del Corral and Antonio Álvarez[♣]

Efficiency Series Paper 07/2004

Abstract: Since it is very likely that all the firms in a sample do not use the same technology it is necessary to consider several reference technologies. In order to control for this unobserved heterogeneity, we estimate these references using two alternative methods. The first one has two stages, where the sample is first split into groups by means of a cluster algorithm and then a technological reference is estimated for each group. The other method, a latent class model, is a single stage method. These two methods are compared with a traditional stochastic frontier which assumes that all firms use the same technology. In the empirical application we estimate a production function using data on Spanish dairy farms.

Key words: dairy farms, latent class models, production function, stochastic frontier.

[*] The authors thank Carlos Arias, Luis Orea and David Roibás for their comments.

[♣] Departamento de Economía, Universidad de Oviedo, Spain. The e-mail address for the corresponding author is: corraljulio@uniovi.es.

1. Introduction

The estimation of production (and cost or profit) functions usually relies on the assumption that the underlying technology is the same for all producers. However, some firms in an industry may use different technologies. In such a case, estimating a common technology to all firms is not appropriate because it can yield biased estimates of the technological characteristics.

To avoid this problem, two-stage processes are sometimes used. In the first step, the sample is split into several groups based on some a priori information about the firms (e.g. private or public ownership, location, etc) or using a cluster algorithm. In the second stage, different functions are estimated for each group. Another alternative is to use models that separate the sample and estimate the technology for each group in only one stage. The latent class models[1] belong to this category. These models classify the sample into several groups, assigning each individual to one group by using the estimated probabilities of class membership.

In this paper we investigate the existence of different technologies in the Spanish dairy sector. The number of dairy farms in Spain has been rapidly decreasing in the last two decades mainly due to the effect of the constraints imposed by the European Union Common Agricultural Policy (e.g., production quotas). Since the country's quota has stayed more or less constant, the remaining dairy farms have increased their size. The fast growth process experienced by many dairy farms has been accomplished by a change in the production system, with some farms adopting more intensive systems. This process of intensification has not been studied properly since intensification is not well defined by only one variable and therefore it is not easy to classify the farms into a specific production system. For this reason, there are not many papers that have studied the technological differences between intensive and extensive dairy farms. The empirical analysis uses data on a balanced panel of 195 Spanish dairy farms over a 5 year period from 1999 to 2003.

We employ a latent class model (e.g., Greene, 2001) to identify several production systems according to their intensification and to estimate different reference

[1] See Greene (2002) for a survey of these models.

technologies. The results obtained with the latent class model are compared with two alternatives. One is the stochastic frontier model proposed by Aigner, Lovell and Schmidt (1977), which assumes that the technology is common to all firms in the sample. In the second alternative, the sample is split into three groups using a cluster algorithm and a production frontier is then estimated separately for each group. Some relevant technological characteristics are compared, namely production elasticities, scale elasticity and marginal products.

The rest of the paper is organized as follows. Section 2 describes the concept of unobservable heterogeneity. Section 3 describes the stochastic frontier latent class model. Section 4 describes the data and the empirical model, while section 5 reports the empirical results. Section 6 analyzes the farms' technical efficiency. Finally, section 7 concludes.

2. Unobservable heterogeneity

In most empirical papers there are differences across observations that are not reflected in the data. This information not observed in the sample is referred to as unobservable heterogeneity. When this information is not important, it can be accommodated in the error term. However, when these differences are important the issue arises as to how to deal with this problem.

Several types of unobservable heterogeneity can be found in practice. The first one is the case when relevant variables are omitted from the model[2] (input quality is a common example). The problem arises when the omitted variables are correlated with some explanatory variables. In such cases the estimated parameters will be biased (Griliches, 1957). A frequent case occurs when the information not included in the model can be considered as invariant over time, e.g., ability in wage models or the soil quality in agricultural production functions. If panel data are available, the solution to this problem is to model the heterogeneity as an individual effect (see Schmidt and Sickles, 1984).

[2] Some reasons for not including a relevant variable are that the variable cannot be measured or its relationship with the dependent variable is unknown to the analyst.

A second type of unobserved heterogeneity refers to the estimation relationship; that is, the firms may come from different data generating processes. This case is shown in figure 1, where there are two groups of firms with different production technologies. In the vertical axis is output, y, while in the horizontal axis is the input quantity, x.

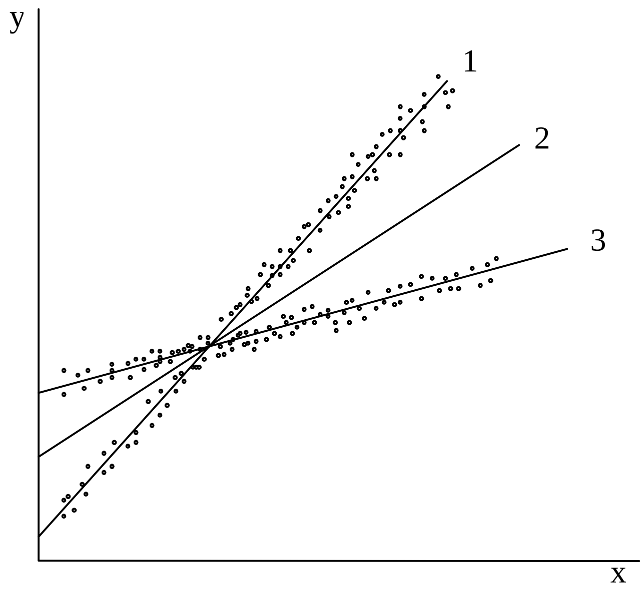

Figure 1. OLS estimation with different technologies

If only one production function is estimated, OLS will fit the function 2, and the estimated technological characteristics (production elasticities, scale economies, marginal products,…) will be biased. To overcome this problem it is necessary to use models that estimate different parameters for each group. Some of these models are continuous in the sense that they allow for a different technology for each firm. The random parameters models (Hildreth and Houck, 1968; Swamy, 1971) or local maximum likelihood (Kumbhakar et. al., 2004) are two examples. Other methods are discrete, in the sense that they create several groups and estimate as many reference technologies as there are groups. This category includes latent class models and cluster analysis. A latent class model (Greene, 2003) assumes that there are a finite number of structures underlying the data. Each firm belongs to one class, though class membership is unknown to the analyst. These models classify the sample into several groups and assign each firm to one group using the estimated class membership probabilities.

On the other hand, the cluster algorithm stratifies the sample in several groups. In a separate step the researcher can estimate a technological reference for each group. This process has two shortcomings. The first is associated with the first stage, since it may happen that the firms are not classified according to their technology. The second one is associated with the efficiency of the estimation, in thatthis procedure does not use the information contained in one class to estimate the technology of firms that belong to other classes. However, in most empirical applications this inter-class information may be quite important because firms belonging to different classes often come from the same industry or sector. Since this kind of information is not exploited, it is possible that two-stage methods are not efficient.

3. Latent class models

We estimate a stochastic production frontier using a latent class model. We begin by briefly describing stochastic frontier models and then the estimation of the latent class model is set out.

Stochastic frontier models were originally proposed by Aigner, Lovell and Schmidt (1977). A stochastic frontier production function may be written as:

$$\ln y = f(x) + \varepsilon \,; \qquad \varepsilon = v - u \quad (1)$$

[3]

where y represents the output of each firm, x is a vector of inputs, f(x) represents the technology, and ε is a composed error term. The symmetric component, v, captures statistical noise and it is assumed to follow a normal distribution with zero mean and standard deviation σ_v. The u term reflects the firm technical inefficiency relative to the stochastic frontier and it is usually assumed to have a half-normal distribution[4], so that $u \geq 0$. Furthermore, the two components v and u are assumed to be independent of each other.

The stochastic frontier framework with the latent class models structure has been used in several recent papers. Caudill (2002) estimates a stochastic cost function with latent

[3] $\varepsilon = v + u$ in cost functions.

[4] Other distributions used in the literature are the exponential, the truncated normal or the gamma.

class structure using US banking data. Orea and Kumbhakar (2004) estimate a stochastic cost frontier using a panel of Spanish banks through a latent class model. Moutinho, Machado and Silva (2003) compare technological change across countries at different stages of development using a latent class model with two groups.

Following the notation in Greene (2001) we can write equation (1) as a latent class model in the following way:

$$\ln y_{it}\big|_j = f(x_{it})\big|_j + v_{it}\big|_j - u_{it}\big|_j \tag{2}$$

where subscript i denotes firm, t indicates time and j represents the different classes (groups). The vertical bar means that there exists a different model for each class j. It is assumed that $v_i\big|_j \sim N(0,\sigma_{vj}^2)$ and $u_i\big|_j \sim \left|N(\sigma_{uj}^2)\right|$.

The Likelihood Function (LF) for each firm at time t can be obtained as a weighted average of its LF from each group j at time t, using as weights the prior probability to membership in class j.

$$LF_{it} = \sum_{j=1}^{J} LF_{ijt} P_{ij} \tag{3}$$

The prior probabilities must be between zero and one: $0 \le P_{ij} \le 1$. Therefore, the sum of these probabilities for each firm must be one: $\sum_j P_{ij} = 1$. In order to satisfy these two conditions we parameterized these probabilities as a multinomial logit. That is:

$$P_{ij} = \frac{\exp(\delta_j q_i)}{\sum_{j=1}^{J} \exp(\delta_j q_i)} \tag{4}$$

where q_i is a vector of variables which are used to divide the sample[5] and δ_j are parameters to be estimated.

The total contribution of each firm to the unconditional likelihood function can be obtained by multiplying the likelihood functions for each period of time considered.

[5] It is also possible not to use separating variables. In this case the latent class model uses the goodness of fit of each estimated frontier as additional information to identify groups of firms.

$$LF_i = \prod_{t=1}^{T} LF_{it} \tag{5}$$

The overall likelihood function is obtained by aggregating the likelihood function of each firm:

$$LF = \sum_{i=1}^{N} LF_i \tag{6}$$

Therefore, the likelihood function can be written as:

$$LF = \sum_{i=1}^{N} \left[\prod_{t=1}^{T} \left(\sum_{j=1}^{J} LF_{ijt} P_{ij} \right) \right] \tag{7}$$

The estimated parameters can be used to estimate the posterior probabilities of class membership, using Bayes Theorem:

$$P(j/i) = \frac{P_{ij} LF_{ijt}}{\sum_{j=1}^{J} P_{ij} LF_{ijt}} \tag{8}$$

An unsolved question in these models is how to determine the number of classes (J is not a parameter to be estimated). According to Greene (2002) testing 'up' from *J*-1 to *J* is not a valid approach because if there are *J* classes, then estimates based only on *J*-1 are inconsistent. However testing 'down' should be valid. Thus, beginning from a *J** known to be at least as large as the true *J*, one can test down from *J** to *J* based on likelihood ratio tests.

Information Criteria such as the Akaike Information Criterion (AIC) or the Schwarz Bayesian Information Criterion (SBIC) are other alternatives to test the number of classes. Moutinho et al. (2003) avail of these in a latent class framework using the following expressions:

$$SBIC = -2 \cdot \ln LF(J) + m \ln(n) \tag{9}$$

$$AIC = -2 \cdot \ln LF(J) + 2 \cdot m \tag{10}$$

where LF(J) is the value that the likelihood function takes for J groups, m is the number of parameters used in the model and n is the number of observations ($n = N \cdot T$, where *N* is the number of firms and *T* is the number of years). The favored model will be that for which the value of the statistic is lowest.

4. Data and empirical model

The data used in the empirical analysis consist of a balanced panel of 195 Spanish dairy farms which were enrolled in a voluntary Record Keeping Program over a 5 year period from 1999 to 2003. Table 1 provides a descriptive summary of the main variables used in the analysis.

Table 1. Data descriptive statistics

	Mean	Standard deviation	Minimun	Maximum
Milk (l.)	290454	0.50	90484	954777
Feed (kg.)	138980	0.56	29777	547487
Cows (units)	39	0.38	20	113
Labor (worker-equivalents)	1.95	0.31	1	4
Land (hectares)	18.91	0.38	10	50
Crop expense (euros)	6302	3698	1493	30587
Feed Cost / Variable cost (%)	74	10	42	92
Milk / Cows	7175	0.18	3586	10973
Milk / Land	15825	0.38	4292	43216
Feed / Cows	3406	0.26	1240	6182
Cows / Land	2.18	0.29	0.91	4.24

The last four rows refer to variables that reflect the intensification of the farm production system. Although the means of the variable show that intensification is relatively important, there exist large differences between the minima and maxima. For instance, there are farms that produce milk with very little land so the majority of nourishment will have to be bought, while other farms base their milk production on their own crop production, so they employ an extensive production system.

The functional form chosen for the production function is the translog. Each explanatory variable in the original data was divided by its geometric mean. In this way, the translog can be considered as an approximation to an unknown function and the first order coefficients can be interpreted as the production elasticities evaluated at the sample geometric mean (point of approximation). The dependent variable is the production of milk (liters)[6]. Four inputs are considered: number of cows (annual average number of dairy cows), labor (full time man-equivalents), feed (annual consumption of feedstuffs in kilograms per farm) and 'crop expense' (which includes the expenses necessary to produce forage crops: seeds and treatments, fertilizers, fuel, machinery hire, diverse materials and both machinery amortization and a land opportunity cost). Additionally, time dummy variables were introduced to control for factors that affect all farms in the same way but vary over time (the period excluded is 1999).

The separating variables employed are 'feed per cow', 'percentage of feed cost' and the 'stocking rate' (number of cows per hectare of land). The variables are measured as ratios in order to avoid farms being grouped by size. Cluster analysis was performed using the means of the separating variables over the five years [7], so that a balanced panel is obtained where each firm is assigned to one group using data from all the years, as in the latent class model[8]. Three groups were obtained: one extensive, composed by 67 farms, a semiextensive group with 89 units and an intensive group integrated by 39 dairy farms.

5. Estimation and results

The latent class model of equation (2) was estimated by maximum likelihood[9]. The model with three groups[10] was the preferred model according to the AIC and the

---

[6] We have considered only one output since these farms are highly specialized (more than 90% of farm income comes from dairy sales).

[7] The cluster analysis has been carried using the k-means algorithm from SPSS 12.0.

[8] Other alternatives would be to estimate a cluster for each year allowing the farms to change system or estimate a cluster for each year assigning each farm to the average system (obtaining a balanced panel).

[9] The estimation was done using LIMDEP 8.0.

[10] Only two separating variables were significant: Feed per Cow and Feed Cost / Variable Cost. 'Feed per cow' has the expected signs, being positive for the intensive group and negative for

’testing down‘ proposed by Greene (2002), while according to SBIC, the model with two groups is the preferred one. The latent class model is compared with the other two models: a stochastic frontier which does not account for heterogeneity; and a model where the sample is split into three groups using a cluster algorithm and a stochastic frontier is then estimated for each group. The estimated parameters of the three models are displayed in the Appendix.

The three models yield different results. These differences are analyzed by comparing the output elasticities (first order coefficients in table A1) and the scale elasticity (the sum of the output elasticities). Table 2 shows the estimated output elasticities of the two main inputs (cows and feed) as well the scale elasticity for all models. In order to make the comparison more homogeneous between the pooled frontier and the other models, the pooled frontier elasticities are evaluated using the groups of the latent class model. There are two kinds of technological differences: differences across models (pooled stochastic frontier, LCM and cluster) and differences across groups within each model (extensive, semiextensive and intensive).

Table 2. Output and scale elasticities

	Cows elasticity			Feed elasticity			Scale elasticity		
	Ext.	Sem.	Int.	Ext.	Sem.	Int.	Ext.	Sem.	Int.
Pooled SF	0.57	0.61	0.62	0.48	0.45	0.44	1.08	1.09	1.13
LCM♠	0.56	0.64	0.68	0.41	0.37	0.39	1.05	1.05	1.13
Cluster	0.57	0.71	0.70	0.40	0.37	0.37	0.97	1.16	1.09

The differences between the intensive and semiextensive groups are quite small in the cluster and in the pooled stochastic frontier model, and a little larger in the latent class model. However, the differences are important between the semiextensive and extensive groups. In all models we can observe that the output elasticity of cows

the extensive group, so a higher value of this variable increases the probability of assigning a farm into a more intensive group.

increases with intensification, while the output elasticity with respect to feed decreases with intensification.

The scale elasticity increases with intensification in all models. This result was unexpected since, in general, more intensive farms are larger and we would expect that when farms grow they exhaust economies of scale. For this reason it was expected that the scale elasticity of the extensive farms would be higher than the scale elasticity of the intensive farms. However, the different farm groups have different technologies, and therefore it is possible that the intensive farms have not exhausted yet their economies of scale. In fact, the results of the cluster model show that the majority of intensive and semiextensive farms (80, 88%) show increasing returns to scale, while this percentage is only (32%) for the extensive group.

The time effects have the same structure in all groups, increasing until 2002 and dropping in year 2003.[11] However, they differ both in significance and in the parameters values, being higher and more significant in the more intensive groups. These time dummy variables basically control two aspects: neutral technological change and the effect of variables that vary over the time but do not vary across farms (e.g. weather).

All the parameter estimates in the pooled stochastic frontier (first order terms, second order terms, cross terms, temporal dummies, λ, σ and scale elasticity), with the exception of the output elasticity of feed, are between the minimum and the maximum from the estimated parameters obtained with the cluster groups. This shows that if unobserved heterogeneity is not controlled for, the estimated parameters are biased.

The marginal products are another important technological characteristic. Table 3 presents the marginal products of two inputs: cows and feed. The values are sample averages over the whole sample period.

[11] 2003 was the driest year, especially from February to May. The experts consider that spring rains are the most important determining factor in forage production.

Table 3. Marginal products (sample average)

	Marginal product of cows			Marginal product of cows		
	Extensive	Semiext.	Intensive	Extensive	Semiext.	Intensive
Pooled SF	3,627	4,462	5,257	0.96	1.01	1.03
LCM	3,371	4,752	5,949	0.86	0.82	0.89
Cluster	3,545	4,848	5,954	0.98	0.79	0.71

There is a clear tendency for the marginal product of cows to increase with intensification. The smallest difference between the intensive and extensive group is 1,630 liters of milk in the pooled stochastic frontier technology while the biggest difference is 2,578 liters of milk in the latent class model. On the other hand, the marginal product of feed does not show a clear tendency. In the cluster model there is a strong tendency to decrease with intensification, while in the latent class model there is no clear trend, with the differences between groups being quite low.

Table 4 shows the evolution of the marginal product of cows and feed over the sample period for all models and groups.

The marginal product of cows follows the same path in all models: it increases until the year 2002 and decreases in 2003. The marginal product of feed also shows a drop in 2003, whereas in the other years it does not vary much.

In Table 5 we check whether the latent class model and the cluster algorithm classify farms into the same groups.

Even though the three groups in each method have almost the same size, they are not composed of the same farms. The largest coincidence occurs in the extensive group (55% of the extensive farms according to the cluster analysis are considered extensive in the latent class model). On the contrary, only 38% of the cluster intensive farms are considered intensive in the latent class model. These large differences show that both methods classify farms in a different manner. In Table 6 we show the differences between the groups formed according to both methods.

Table 4. Evolution of marginal products

Model	Group	Input	1999	2000	2001	2002	2003
Pooled Frontier		Cows	4,096	4,200	4,354	4,482	4,377
		Feed	1.01	1.01	1.00	1.03	0.96
Latent Class	Extensive	Cows	3,215	3,274	3,404	3,577	3,387
		Feed	0.91	0.88	0.85	0.88	0.79
	Semiextensive	Cows	4,519	4,611	4,798	4,938	4,893
		Feed	0.82	0.82	0.82	0.84	0.80
	Intensive	Cows	5,668	5,945	6,074	6,077	5,983
		Feed	0.83	0.89	0.91	0.92	0.90
Cluster	Extensive	Cows	3,377	3,411	3,603	3,736	3,597
		Feed	1.02	1.00	0.98	1.00	0.90
	Semiextensive	Cows	4,975	5,131	5,293	5,390	5,276
		Feed	0.78	0.79	0.79	0.80	0.77
	Intensive	Cows	5,678	5,864	5,996	6,251	6,182
		Feed	0.71	0.69	0.71	0.75	0.71

Table 5. Cluster and LCM group classification

		Cluster Analysis			
		Extensive	Semiextensive	Intensive	Total
Latent Class Model	Extensive	37	24	7	68
	Semiextensive	29	49	17	95
	Intensive	1	16	15	32
	Total	67	89	39	195

Table 6. Descriptive statistics of the groups

	Latent class model			Cluster analysis		
	Ext.	Semiext.	Intensive	Ext.	Semiext.	Intensive
Observations	340	475	160	335	445	195
Milk liters	225,778	278,752	462,630	210,169	285,384	439,949
Cows	37	37	52	34	38	51
Land	19	18	23	18	18	23
Labor	1.92	1.85	2.31	1.75	1.98	2.24
Milk per Cow	6,050	7,442	8,768	6,144	7,351	8,543
Milk per Land	12,731	16,400	20,688	12,534	16,289	20,417
Milk per Feed	2.10	2.21	2.25	2.42	2.11	1.90
Feed per Cow	3,022	3,479	4,002	2,589	3,520	4,546
Cows per Land	2.09	2.18	2.34	2.03	2.20	2.37
Feed Cost (%)	75	73	72	72	74	75

In general, the characteristics are similar in both methods: more intensive farms produce more milk, use more inputs, have high-yielding cows, consume more feed per cow, have more cows per hectare and, as a result, they produce more milk per hectare. The main difference appears in the feed average product, which is greater in the intensive group of the latent class model and greater in the extensive regime in the cluster classification. Another difference appears in 'Feed cost %' which in the latent class model decreases with intensification wheres in the cluster groups it grows with intensification. These differences can be explained by the fact that the cluster analysis splits the sample according to the values of the separating variables while the latent class model uses the effects of the separating variables on the dependent variable.

We carry out some tests of average differences. On the one hand we compare the differences between the groups labeled with the same name, and on the other we compare the differences between the semiextensive group with the intensive and the extensive ones from the same method. The test results show that there are significant differences between groups formed in each method, while the evidence with respect to the groups labeled with the same name is not very conclusive, with the evidence tending towards indicating that the groups are not different.

Therefore, even though both models use the same separating variables, the classification differs according group characteristics. This suggests that the latent class model outperforms the cluster algorithm due to the fact that the latent class model considers a multinomial logit and also the goodness of fit of the different production functions in order to assign the individuals to each group, whereas the cluster algorithm simply maximizes the inter-group variance and minimizes the intra-group variance.

6. Technical Efficiency

Technical efficiency (TE) reflects the ability of a farm to produce the maximum level of output from a given set of inputs. In this section we estimate farm-specific TE indexes for the three models. In the stochastic frontier model an output-oriented technical efficiency index is calculated using the following expression:

$$TE_{it} = \exp(-\hat{u}_{it}) \qquad (11)$$

Even though our data set forms a balanced panel, the stochastic frontiers estimated in the three models are "pooled" in the sense that they impose independence over time of the U_{it}. This implies that each observation is treated as a different farm and therefore the statistical analysis does not exploit the feature that the same farms are observed repeatedly over time.

In the latent class model the calculation of the technical efficiency indexes is not so immediate because each farm has several reference frontiers with an associated probability. Two alternative solutions have been proposed in the literature: the reference technology can be the most likely frontier (the one with the highest posterior probability), or the efficiency indexes can be calculated as a weighted average of the technical efficiencies for all the reference technologies using the posterior probabilities as weights:

$$\ln TE_{it} = \sum_{j=1}^{J} P(j/i) \cdot \ln TE_{it}\big|_{j} \qquad (11)$$

The difference between these two alternatives will be higher when the highest posterior probability is lower. We have chosen the second alternative since this contains all the information given by the latent class model while the first alternative omits part of the posterior probabilities, which is information that could be relevant.

Table 7 shows the statistics of the technical efficiency indexes obtained in each model. The highest mean technical efficiency is obtained in the latent class model (94%) while in the other models the average technical efficiency is around 80%. This result was expected because the latent class model has a higher capacity to distinguish between technological differences and can thus better separate the farms that use different technologies. Therefore, it does not attribute the differences in the technology employed by the farms to inefficiency.

Table 7. Technical efficiency indexes

	Mean			Minimum			Maximum		
	Ext.	Sem.	Int.	Ext.	Sem.	Int.	Ext.	Sem.	Int.
Pooled SF	0.81	0.91	0.93	0.47	0.67	0.74	0.96	0.97	0.98
Latent CM	0.94	0.97	0.96	0.75	0.87	0.88	0.99	0.99	0.99
Cluster	0.86	0.88	0.89	0.46	0.48	0.61	0.98	0.98	0.99

7. Conclusions

In this paper we have estimated a production function using three different methods: a pooled stochastic frontier, a stochastic frontier latent class model and separate stochastic frontiers for groups obtained from a cluster analysis. These methods yield different results. In particular we have shown that output elasticities, marginal products and scale economies are different, not only across groups but also between methods.

From a policy point of view, an important result is that the intensive and semiextensive dairy farm scan still increase size to exhaust their economies of scale, whereas it seems that the extensive farms are close to their optimal scale.

References

Aigner, D., K. Lovell and P. Schmidt, (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics,* 6, 21-37.

Battese, G. y T. Coelli, (1992), "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", *Journal of Productivity Analysis,* 3, 153-169.

Caudill, S., (2002), "Estimating a Mixture of Stochastic Frontier Regression Models via the EM: A Multiproduct Cost Function Application", *Empirical Economics,* 28(3), 581-598.

Greene, W., (2001), "New Developments in the Estimation of Stochastic Frontier Models with Panel Data", Efficiency Series Paper 6/2001, Dpto. de Economía, Univ. de Oviedo.

Greene, W., (2002), "Alternative Panel Data Estimators for Stochastic Frontier Models", Working Paper, Dept. of Economics, Stern School of Business, NYU.

Greene, W., (2003), "Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization´s Panel Data on National Health Care Systems", Working Paper, Dept. of Economics, Stern School of Business, NYU.

Griliches, Z., (1957), "Specification Bias in Estimates of Production Function*", Journal of Farm Economics*, 39, 1, 8-20.

Hildreth, C. and J. Houck, (1957), "Some Estimators for a Linear Model with Random Coefficients", *Journal of the American Statistical Association,* 63, 584-595.

Jondrow, J., K. Lovell, I. Materov and P. Schmidt, (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model", *Journal of Econometrics,* 19, 233-238.

Kumbhakar, S. and E. Tsionas, (2003), "Recent Developments in Stochastic Frontier Modelling", Efficiency Series Paper 6/2003, Dpto. de Economía. Univ. de Oviedo.

Kumbhakar, S., E. Tsionas, B. Park, L. Simar, (2004), "Nonparametric Stochastic Frontiers: A Local Maximum Likelihood Approach", Working Paper Series 2004, 0405, University of Binghamton.

Moutinho, N., F. Machado and E. Silva, (2003), "An International Comparison of Productivity Change in Agriculture and the Economy as a Whole", Discussion Papers 9-2003, Facultade de Economia, Universidade de Porto.

Orea, L. and S. Kumbhakar, (2004), "Efficiency Measurement using a Stochastic Frontier Latent Class Model", *Empirical Economics*, 29, 169-183.

Schmidt, P. and R. Sickles, (1984), "Production Frontier and Panel Data", *Journal of Business and Economics Statistics,* 2, 367-374.

Swamy, P., (1971), "Statistical Inference in Random Coefficient Regression Models". Springer, New York.

APPENDIX

Table A1. Estimation of the Latent class model with three groups

	Intensive Group	Semiextensive Group	Extensive Group
Constant	12.63***	12.55***	12.45***
Cows	0.678***	0.638***	0.557***
Feed	0.394***	0.370***	0.411***
Labor	0.022	0.030	0.056*
Crop expenses	0.036*	0.017	0.029**
0.5 · Cows · Cows	0.115	-0.015	-0.005
0.5 · Feed · Feed	-0.110	0.313***	0.280**
0.5 · Labor · Labor	0.000	-0.073	-0.345***
0.5 · Crop$ · Crop$	0.084*	-0.046	0.055
Cows · Feed	-0.106	-0.271*	-0.264
Cows · Labor	0.644***	0.045	0.461***
Cows · Crop$	-0.147	0.026	0.123
Feed · Labor	-0.446***	0.048	-0.157**
Feed · Crop$	0.062	0.041	-0.075*
Labor · Crop$	-0.012	0.001	-0.094**
D2000	0.058***	0.006	-0.008
D2001	0.072***	0.036***	0.009
D2002	0.073***	0.056***	0.052***
D2003	0.056***	0.033***	-0.009
Constant	-3.84		-3.65
Cows/Land	0.038		0.098
Feed/Cows	0.001***		-0.001***
Feed cost %	-0.841		9.191***
$\lambda = \sigma_u / \sigma_v$	2.10	1.63	2.55
$\sigma = [\sigma_v^2 + \sigma_u^2]^{1/2}$	0.09	0.09	0.14
Likelihood Function		1049	
Observations		975	
Scale elasticity	1.13	1.06	1.05

* ,**,*** significant at the 1%, 5% and 10% level respectively

Table A2. Stochastic frontier estimations for the alternative models

	Stochastic Frontier "Pooled"	Cluster groups stochastic frontiers		
		Intensive Group	Semiextensive Group	Extensive Group
Constant	12.60***	13.02***	12.59***	12.38***
Cows	0.600***	0.702***	0.709***	0.577***
Feed	0.459***	0.375***	0.372***	0.405***
Labor	0.026*	0.012	0.063**	-0.030
Crop$	0.012	0.004	0.017	0.022
0.5 · Cows · Cows	-0.544***	0.771	-0.887	-0.873
0.5 · Feed · Feed	-0.233**	0.834*	-0.580	-0.820***
0.5 · Labor · Labor	-0.067	0.006	0.045	-0.277*
0.5 · Crop$ · Crop$	-0.052*	-0.083	-0.021	-0.078
Cows · Feed	0.293**	-0.930*	0.534	0.729**
Cows · Labor	0.161**	0.448*	0.273	0.012
Cows · Crop$	0.033	0.488	-0.025	0.068
Feed · Labor	0.013	-0.408*	0.021	0.070
Feed · Crop$	0.008	0.069	0.116	0.029
Labor · Crop$	-0.014	-0.126*	-0.047	0.016
D2000	0.012	0.008	0.021	-0.006
D2001	0.026**	0.026	0.028	0.021
D2002	0.046***	0.051*	0.032*	0.060***
D2003	0.011	0.033	0.015	-0.002
$\lambda = \sigma_u / \sigma_v$	2.27	2.83	2.07	3.97
$\sigma = [\sigma_v^2 + \sigma_u^2]^{1/2}$	0.17	0.15	0.16	0.19
Likelihood Function	738	178	350	261
Observations	975	195	445	335
Scale elasticity	1.10	1.09	1.16	0.97

* ,**,*** significant at the 1%, 5% and 10% level respectively

Table A3. Mean difference test

	Milk per Cow	Feed per Cow	Cow per Land
Int. Cluster vs int. LCM	YES	YES	NO
Sem. Cluster vs sem. LCM	NO	NO	NO
Ext. Cluster vs ext. LCM	NO	YES	NO
Int vs sem. LCM	YES	YES	YES
Int vs sem. cluster	YES	YES	YES
Ext. vs sem. LCM	YES	YES	YES
Ext vs sem. cluster	YES	YES	YES