universidad de león

Department of Economics and Statistics

Departamento de Economía y Estadística

Working Papers
Documentos de Trabajo

Effects of Switching between Production Systems in Dairy Farming

Antonio Alvarez and Carlos Arias

WP 2014/06

Effects of Switching between Production Systems in Dairy Farming

Antonio Alvarez[1] and Carlos Arias[2]

Abstract

The recent trend in the intensification of dairy farming in Europe has sparked an interest in studying the economic consequences of this process. However, classifying empirically farms as extensive or intensive is not a straightforward task. In recent papers, *Latent Class Models* (*LCM*) have been used to avoid an ad-hoc split of the sample into intensive and extensive dairy farms. A limitation of current specifications of *LCM* is that they do not allow farms to switch between different productive systems over time. This feature of the model is at odds with the process of intensification of the European dairy industry in past decades. We estimate a single LCM that allows for changes of production system over time by estimating a single *LCM* model but splitting the original panel into two periods and find that the probability of using the intensive technology increases over time. Our estimation proposal opens up the possibility of studying the effects of intensification not only across farms but also over time.

JEL: C23, D24, Q12

Key words:

Dairy farms, intensification, latent class model, panel data

[1] University of Oviedo. alvarez@uniovi.es

[2] Corresponding autor. University of León and *Oviedo Efficiency Group*. carlos.arias@unileon.es

1. Introduction

The number of dairy farms in the European Union has fallen dramatically in recent decades and continues to decline. Given that farm output remains roughly at quota level, over the same period the average size of dairy herds has increased steadily. At the same time, genetic and management improvements in dairy cattle have permitted large increases in milk production per cow. These structural changes have provided the basis for the propagation of intensive systems of production in the dairy sector.

Extensive dairy farming, on the other hand, consists of producing milk using mainly on-farm produced forage with low stocking rates. Fostering production using extensive systems has often been an explicit goal of agricultural policy, justified by factors such as environmental soundness, improved animal welfare, or use of abundant land in some areas. In contrast to this, many dairy farms in Europe have gone in the opposite direction, adopting more intensive production systems.

Despite the importance of the intensification process and the intense policy debates it has generated, few papers have studied the intensification of dairy farming using economic analysis. Most articles have adopted a technical perspective, describing the physical changes of the process (e.g., Simpson and Conrad, 1993), while very few have analyzed the economic consequences of this process (some exceptions are Alvarez *et al*., 2010, and Nehring *et al*., 2011).

From an empirical point of view, the coexistence of both extensive and intensive farms implies that there are two different technologies in the sector. This runs contrary to the assumption of a common technology for all farms which is the most frequent in the production literature. However, there is also awareness among researchers of the estimation bias that arises if such an assumption is unrealistic. For this reason, several approaches have been followed in dairy sector studies in order to account for the likely existence of different technologies. The most basic one is to drop a number of farms from the sample on the grounds that they may operate under a different technology (Tauer and Belbase, 1987). A second approach is to split the sample into several groups based on some observable farm characteristics. For example, Hoch (1962) divided the sample into two groups based on the location of farms, while Newman and Matthews (2006) consider two different technologies depending on the number of outputs each farm produces.

Classifying farms as intensive or extensive is not as straightforward as it might appear. For example, Nehring et al. (2011) used the number of cows per hectare in order to split the sample. However, the stocking rate partition does not fully describe the production system. Other aspects such as the productivity of cows or the share of concentrates in the feed ration could also be taken into account. In this paper we avoid an ex-ante classification of farms as extensive or intensive by estimating a Latent Class Model (*LCM*). This model assumes that several unknown technologies (classes) have generated the sample, and allows for the estimation of the parameters of the different technologies plus the probability that each observation has been generated by a specific technology.

To the best of our knowledge, the starting point of previous models in the literature is a set-up and estimation proposal that assumes that the probability of each observation belonging to a class (i.e. using an intensive or extensive technology) is constant over time. This implies that changes during the period of analysis, no matter how dramatic, do not lead to a farm being labeled as belonging to a different class (use of a different technology). Such an assumption becomes increasingly untenable as the number of observed periods gets larger.

The objective of the present paper is twofold. First, we wish to determine whether the intensification process that has been taking place in recent decades has come to an end or whether dairy farms are still switching from extensive to more intensive production systems. For this purpose, we make a simple methodological proposal to circumvent the assumption of class probability being constant over time. Second, we are interested in analyzing the effects of intensification on farms' efficiency. In particular, for given inputs we would like to know whether intensive farms have the potential to produce more output than extensive farms, and if so, the degree to which intensive farms achieve such potential.

In order to fulfill these two objectives, in the empirical section of the paper, we use a panel of dairy farms in Northern Spain to illustrate the feasibility of estimating a *LCM* with time varying probabilities. Our objectives are first to measure the changes over time in the probability of belonging to a class (technology) and second to analyze the effects on production potential and efficiency of dairy farm intensification[3]. Unlike previous studies, our methodological proposal allows us to look specifically at the farms that might have tilted towards a more intensive production system in the period of analysis.

The organization of the paper is the following. In next section we describe the model. In section 3, we present the data and the empirical model. In section 4 we show the econometric estimation of the latent class models. In section 5 there is a discussion of the empirical results. The paper ends with some conclusions.

2. The Model

We use a *LCM* to analyze the extent of intensification in our dataset. The initial step is to check whether the *LCM* identifies different technologies and if those technologies represent different degrees of intensification. The starting point is a log-linear stochastic production frontier (Orea and Kumbhakar, 2004) such as:

$$ln\, y_{it} = ln\, f\left(x_{it}\right)|_j + v_{it}\,|_j - u_{it}\,|_j \qquad (1)$$

where x is a vector of inputs, y a single output, v a symmetric random disturbance, and u a one-sided random disturbance that measures technical inefficiency. Subscript i $(i=1,\ldots,N)$ denotes firms, t $(t=1,\ldots,T)$ denotes time, and subscript j $(j=1,\ldots,J)$ indicates a technology in a finite set. The vertical bar means that there is a different production function (different parameters) for each class j. We assume that, conditional on each class, the random disturbances v and u follow a normal and a truncated normal random distribution respectively.

In a latent class model we need to consider three likelihood functions. The first is the likelihood function of a firm i at time t belonging to class j:

$$LF_{ijt} = g\left(y_{it}, x_{it}, \Theta_j\right) \qquad (2)$$

[3] The present paper is related with a strand of literature linking technological choices with efficiency. As an example, Kompas and Nhu Che (2006) studied the effect of different technologies on the efficiency of dairy farms by including in the inefficiency model a set of variables reflecting technological choices.

where Θ_j represents the set of parameters of technology (class) j, and g denotes the likelihood function of a production frontier (Kumbhakar and Lovell, 2000). The second is the likelihood function of a firm i conditional on class j, obtained as the product of the likelihood functions in each period.

$$LF_{ij} = \prod_{t=1}^{T} LF_{ijt} = \prod_{t=1}^{T} g\left(y_{it}, x_{it}, \Theta_j\right) \tag{3}$$

Finally, the unconditional likelihood function of firm i is calculated averaging the likelihood conditional on each class using the prior probabilities of class membership P_{ij} as weights:

$$LF_i = \sum_{j=1}^{J} LF_{ij} P_{ij} \tag{4}$$

Prior probabilities can be interpreted as the probabilities attached to membership of class j (Greene, 2005). These prior probabilities can be parameterized using a multinomial logit model such as:

$$P_{ij} = \frac{exp\left(\delta_j z_i\right)}{\sum_{j=1}^{J} exp\left(\delta_j z_i\right)} \tag{5}$$

where z_i is a vector of "separating variables" and δ_j a vector of parameters to be estimated. The "separating variables" are related to the adoption of a technology and, as a result, can be used as explanatory variables of the prior probability of using that technology.

After estimation of the model in (1) by maximum likelihood a "posterior probability" can be computed as:

$$Pr_{ij} = \frac{LF_{ij} P_{ij}}{\sum_{j=1}^{J} LF_{ij} P_{ij}} \tag{6}$$

As equation (6) shows, the prior probability of class membership for each farm is weighted by the empirical likelihood that the farm belongs to that class. This implies that the ability of each technology to explain the observed production of a farm is incorporated in the calculation of the posterior probability. In a sense, the estimates obtained with the parsimonious parametric model of the prior probability are complemented with information on individual fit to provide a more accurate evaluation of the probability of class membership. In fact, the posterior probability is considered the best estimate of class membership (Greene, 2005), and as such, the value of this probability is the criterion that we will use for determining whether a farm is using a particular technology.

As mentioned above, a subtle feature of the *LCM* for panel data is that prior probabilities are modeled as time invariant. In practical terms, time-invariant probabilities amount to assuming that changes in farms over time don't affect the adscription of a farm to a class (intensive or extensive technology). However, we expect some farms to change the use of inputs during the period of analysis in ways that suggest tilting towards intensive farming.

Our aim, therefore, is to circumvent the assumption of time-invariant probabilities in the empirical analysis. For that purpose, we estimate a single *LCM* model for the whole period of analysis but split the observations for each farm into two periods, where the first period roughly corresponds to ($t=1,\ldots,T/2$) and the second period to ($t=T/2+1,...,T$). In this approach, farm i is considered to be a different farm in each of the two periods. As a result, the probability of class membership is constant for a given farm during each of the two periods but can change from the first period to the second.

An alternative approach to obtain time-varying probabilities would be to estimate a pooled model. This amounts to considering the dataset as a cross section of farms instead of taking into account the existence of a panel. In this case, in each observation farm i is treated as a different farm, thereby allowing probabilities of class membership to vary freely over time. The downside is that we disregard the information provided by observing the same farm in several periods and that the computed probabilities of class membership are not necessarily parsimonious. Indeed, it would be possible to observe some farms moving repeatedly from one class to the other. All in all, this approach seems prone to numerical problems and to difficulties for interpreting the results. In fact, in the empirical application that we propose the *LCM* with pooled data failed to converge.

At this point, we should note that Alvarez and del Corral (2010) report time-varying probabilities of class membership in the conventional Panel LCM. They get that result through a slight modification of equation (6). They use as weighting factor of prior probabilities the likelihood function of each observation in each year (LF_{ijt} in equation 2) instead of the likelihood function of each farm (LF_{ij} in equation 3). However, they report time variation of a feature of the model, latent class probability, assumed to be constant over time for estimation purposes. Two things can happen. First, if the latent class probability is indeed time-varying the model used for estimation is misspecified. Second, if the latent class probability is constant over time the model is correctly specified but you are reporting time variation akin to variations around a mean. Our estimation proposal avoids such problems by allowing class probability to change from the first subperiod to the second while being constant in each subperiod.

3. Data and empirical model

The data used in the empirical analysis consist of a balanced panel of *128* Spanish dairy farms observed over the *12* year period from *1999* to *2010*. In general, the units considered are small to medium-sized family farms.

The empirical specification of the production function is a translog. We have chosen a flexible functional form in order to avoid imposing unnecessary a priori restrictions on the technologies to be estimated. The empirical counterpart of equation (1) is the following translog production frontier:

$$ln\, y_{it} = \beta_0 |_j + \sum_{k=1}^{5} \beta_k |_j \, lnx_{kit} + \frac{1}{2} \sum_{k=1}^{5} \sum_{l=1}^{5} \beta_{kl} |_j \, ln\, x_{kit}\, lnx_{lit} + \sum_{m=2}^{12} \gamma_m |_j \, TD_m + v_{it} |_j - u_{it} |_j \qquad (7)$$

The dependent variable (y) is the production of milk (liters). We have considered only one output since these farms are highly specialized (more than 90% of farm income comes from dairy sales). Five inputs are included: (x_1) number of cows, (x_2) purchased feed (kilograms), (x_3) ‘farm expenses’ (includes expenditure on inputs used to produce forage crops, namely seeds, sprays, fertilizers, fuel, and machinery depreciation), (x_4) ‘animal expenses’, such as veterinary, medicines, milking and other expenses, and (x_5) land. All monetary variables are expressed in constant euros of 2004. Additionally, 11 time dummy variables ($TD_m=1$ *if* $t=m$, $TD_m=0$ *otherwise*) were introduced to control for factors that affect all farms in the same way each year but which vary over time, such as weather (the excluded period is 1999).

Prior to estimation each input was divided by its geometric mean. In this way, the first order coefficients of the Translog production function (β_k^j) can be interpreted as output elasticities evaluated at the geometric mean of the inputs.

The prior probabilities of class membership are assumed to be a function of two "separating variables": the natural logarithm of the stocking rate (cows per hectare) and the natural logarithm of concentrate feed per cow[4].

4. **Econometric estimation**

In this section, we report the main results of the estimation of two latent class models by maximum likelihood:

a) Panel model, i.e. using panel data and time-invariant class membership probabilities.

b) Split-panel model, i.e., estimating a single *LCM* model but allowing the probabilities of class membership to differ over time by splitting the sample into two periods where the probability of class membership is constant within each period but can change for the 'same' farm from the first period to the second. Precisely, the estimation proceeds by treating each farm as a different farm in the second period.

In both models two latent classes were found.[5] As mentioned above, we make the prior probability of belonging to a latent class a function of two variables: 'cows per hectare of land' and 'feed per cow'. Since these variables measure the degree of intensification of a dairy operation, we have labeled as "intensive" the latent class which shows positive estimates of the coefficients of both 'separating' variables in the prior probability equation. The estimates of the prior probability function of the intensive class for both models are shown in Table 1.

---

[4] Since prior probabilities are modeled as time-invariant for a given period of time, the explanatory variables are averaged over such period.

[5] We also tried to fit a model with three classes but it did not converge.

Table 1: Prior probability equation for the intensive class

	Panel model	***Split-panel model***
Constant	*-11.906**	*-13.112***
ln(cows/land)	*1.7919***	*.8632**
ln(concentrate/cows)	*1.3510**	*1.5513***

**, **, Significantly different from zero at 0.05 and 0.01 significance levels respectively*

Standard errors reported in Tables A1 and A2 in the Appendix

Next, the farms were classified as *Intensive* using the highest (greater than 0.5) estimated posterior probabilities (equation 6) since these provide the best estimates of class membership for an individual (Greene, 2005). In Table 2 we show descriptive statistics of the two groups (intensive and extensive) for the two models estimated.

Table 2. Characteristics of the estimated production classes (sample means)

	Panel model		***Split-panel model***	
	Intensive	***Extensive***	***Intensive***	***Extensive***
Observations	*804*	*732*	*798*	*738*
Milk (l)	*365442*	*280884*	*371525*	*274994*
Cows	*42.7*	*40.3*	*43.2*	*39.9*
Land (ha)	*18.0*	*19.8*	*18.3*	*19.6*
Cows per hectare	*2.45*	*2.16*	*2.4*	*2.1*
Milk per cow (l)	*8129*	*6745*	*8202*	*6677*
Milk per hectare (l)	*20329*	*14779*	*20428*	*14718*
Feed per cow (Kg)	*3533*	*3352*	*3579*	*3304*

The descriptive statistics of each group roughly agree with the labels we gave to the classes based on the effects of intensification variables on the probability of being in each class.

As expected, intensive farms have larger values of key variables such as milk per cow, milk per hectare and feed per cow. Intensive farms are also larger in terms of milk production but are rather similar in terms of land. In our view, the explanation for this result is that marginal increases of land are unlikely to be an option for farmers due to the fact that most abandonments take place in less favored areas (mountainous) while remaining farms are mainly located in the coastal plain, so that the land available after some farms shut down cannot be used by the remaining farms. For this reason, farmers who wish to increase production need to use more feed per cow and in some cases buy more productive cows, thereby becoming more intensive.

In Table 3 we report the output elasticities for the two groups evaluated at the geometric mean of the sample. The differences in the elasticities across groups can be seen as evidence of different technological characteristics. The complete set of estimated parameters of the production functions are reported in Tables *A1* and *A2* in the Appendix.

Table 3. Output elasticities evaluated at the geometric mean of the sample

	Panel model		***Split-panel model***	
	Intensive	***Extensive***	***Intensive***	***Extensive***
Cows	*.7176***	*.4626***	*.7491***	*.4553***
Feed	*.2969***	*.3623***	*.2737***	*.3685***
Farm expenses	*.0405***	*.0718***	*.0508 ***	*.0767***
Animal expenses	*.0296***	*.0964***	*.0206 **	*.0789***
Land	*.0368***	*.0339***	*.0214*	*.0295**

**, **, Significantly different from zero at 0.05 and 0.01significance level, respectively*

All elasticities, with one exception, are significantly different from zero at conventional levels of significance. Despite the different assumptions behind the two models the output elasticities evaluated at the sample geometric mean are similar across models. However, it is interesting to note that there are wide differences between the parameters of the two latent classes (within each model). For example, the output elasticity with respect to cows is almost twice as large in the intensive group as it is in the extensive group. On the other hand, the output elasticity of feed is always larger in the extensive group. These different elasticities imply large differences in marginal productivity of inputs across technologies (extensive or intensive), especially for cows and feed.

5. Empirical extensions

In this section we use the results of the estimation of the *LCM* to analyze a set of issues with important policy implications. First, we are interested in studying the evolution of the intensification process over time. Second, we want to analyze the differences in technical efficiency between intensive and extensive farms.

5.1. Evolution of intensification over time

We want to check if the probability of adopting the intensive technology increases over time. We should note that this analysis is only possible in the split-panel model and not in the conventional panel *LCM*.

For this purpose, we regress the posterior probabilities of being in the intensive class against individual dummies (fixed effects) and a binary variable (d_t) that takes the value zero for the first half of the period analyzed and one for the second half. This is an unconditional analysis over time. It is clearly different from the conditional analysis of prior probabilities that could be achieved by including a time trend in equation (5). In the conditional analysis, prior probabilities could change over time, keeping input use and separating variables constant. In the unconditional analysis performed here, the posterior probability varies over time due to changes in the use of key inputs such as feed, cows or land.

The equation to be estimated is the following:

$$Pr_{ijt} = a_i\,|_j + b\,|_j\, d_t + w_{it}\,|_j \qquad (8)$$

where *j* denotes the latent class and *w* is a random disturbance. For each class, we have an estimated posterior probability for each individual (*i=1,…,128*) and for each period (*t=1,..,12*). Expression (8) represents a general proposal with *J* different classes. In this case, there would only be *J-1* free equations since the dependent variable, the posterior probability, adds up to one. As we are considering only two latent classes (Intensive and Extensive) in our setting, we have only one relevant equation. We choose to estimate the equation corresponding to the posterior probability of the *Intensive* Class.

Table 4. Analysis of the evolution over time of posterior probabilities

Coefficient ($b|_{intensive}$)	***Standard Error***	***t-statistic***
.0363	*.0099*	*3.28*

Table 4 shows the estimated coefficient of the time binary variable ($b|_j$) for the Posterior probability of *Intensive Class* in the Split-panel model. This coefficient is positive and significantly different from zero indicating that the probability of being classified as an intensive farm is larger in the second period. We interpret this result as evidence of “intensification” of dairy production in our sample over the period analyzed.

In our view, the average change over time of the probability of belonging to the intensive class provides evidence of farms in the sample tilting towards an intensive production system. Additionally, the change in class probability between the two periods allows for some descriptive analysis of the subset of farms that have switched production system in the conventional sense of crossing the threshold defined by a class probability of 0.5. In particular, 13 extensive farms became intensive in the second period, while 8 farms switched from intensive to extensive. In table 5 we show the characteristics of the farms that change to a different production system.

Table 5. Characteristics of the farms that switch production system over time

	From extensive to intensive		***From intensive to extensive***	
	Period 1 Extensive	***Period 2 Intensive***	***Period 1 Intensive***	***Period 2 Extensive***
Farms	*13*	*13*	*8*	*8*
Milk (l)	279777	383108	269831	301237
Cows	37.2	42.9	35.3	42.5
Land (ha)	16.5	16.6	14.8	17.7
Cows per hectare	2.1	2.4	2.4	2.4
Milk per cow (l)	7061	8381	7456	7004
Milk per hectare (l)	15796	21242	18246	17191
Feed per cow (Kg)	3483	3824	3307	3204

The farms that become intensive in the second period reflect the typical transformation pattern: increase in the stocking rate and feed per cow, resulting in higher milk per cow and per hectare. On the other hand, the farms that switch from the intensive to extensive class keep the stocking rate unchanged but reduce the amount of feed per cow, lowering milk per cow and per hectare.

5.2. The effect of intensification on dairy farm efficiency

In this section, we want to explore the effect of intensification on production efficiency. Two questions are addressed. First, are intensive farms more efficient than extensive farms? Second, which technology is more productive (i.e., does one of the two frontiers lie above the other one)?

The level of technical efficiency can be calculated as the ratio between current output and potential output, as defined by the technological frontier. An output-oriented index of technical efficiency can be computed as:

$$TE_{it}\,|_j = exp\left(-u_{it}\,|_j\right) \tag{9}$$

Subscript *j* in equation (9) indicates that the technical efficiency index can be calculated with respect to each of the Latent Class frontiers (Orea and Kumbhakar, 2004). In our case, we can thus consider two different frontiers. Table *6* shows the average technical efficiency for the two technologies.

Table 6. Average efficiency by model and Latent Class (intensive/extensive)

	Panel Model		***Split-Panel Model***	
	Intensive Frontier	***Extensive Frontier***	***Intensive Frontier***	***Extensive Frontier***
Full sample	*.92*	*.92*	*.91*	*.93*
Intensive farms	*.94*	*.96*	*.94*	*.96*
Extensive farms	*.88*	*.89*	*.87*	*.89*

We would like to point out two results. First, for the full sample, the average level of technical efficiency is quite similar both across models (Panel vs. Split-Panel) and across latent technologies (Intensive frontier vs. Extensive frontier). However, if we consider the two groups of farms separately, a very interesting result is found: intensive farms have higher level of technical efficiency in all the four frontiers considered. Additionally, the average technical efficiency is higher in the extensive frontier than in the intensive one.

This last result seems to indicate that the latent frontier of the *Intensive Group* dominates the other, that is, for any given set of inputs it is possible to produce more output with the intensive technology. We try to shed some light on this issue by calculating the difference in frontier output between the frontiers using the actual inputs of the farms:

$$D_{it} = ln\,\hat{y}_{it}^{I} - ln\,\hat{y}_{it}^{E} \tag{8}$$

where $ln\,\hat{y}_{it}^{I}$ ($ln\,\hat{y}_{it}^{E}$) is the (log of) frontier output of farm *i* at time *t* evaluated at the intensive (extensive) technology. Table *7* shows the average value of D_{it} for the full sample as well as for the two classes.

Table 7: Average difference between intensive and extensive production frontiers

	Panel model	*Split-panel model*
Full sample	*.0962*	*.1080*
Intensive farms	*.0883*	*.0988*
Extensive farms	*.1049*	*.1180*

As the differences between the frontiers are calculated using natural logs, D_{it} can be interpreted approximately as the percentage difference of potential output between both frontiers. For the panel model, the intensive frontier is, on average, *9.6%* above the extensive frontier. For the split-panel model, the intensive frontier is, on average, *10.8*% above the extensive frontier.

Additionally, two interesting issues can be studied: the evolution of technical efficiency over time and the relationship of technical efficiency with the probability of being in a latent class. These two issues can be analyzed both in the conventional panel model and the Split-panel model proposed in the present paper. However, in the Split-Panel model the probability of class membership can change over time. This feature suggests the need for a joint analysis of the effects of time and probability of class membership.

Table 8. Relationship of technical efficiency with intensification

	Panel model		***Split-panel model***			
	OLS		***OLS***		***OLS with farm dummies***	
	Intensive frontier	***Extensive frontier***	***Intensive frontier***	***Extensive frontier***	***Intensive frontier***	***Extensive frontier***
Probability of intensive class	*0.0607***	*0.0658***	*0.0786***	*0.0752***	*0.0545***	*0.0439***
Time dummy	*-0.0010***	*0.0009***	*-0.0014***	*0.0007**	*-0.0013***	*0.0008***

**, **, Significantly different from zero at 0.05 and 0.01significance level, respectively*

In Table 8, we show the results of regressing the level of technical efficiency against the probability of being in the intensive class and a time dummy that takes the value 1 for the second period. We show the results for both models (Panel and Split-panel) and for both frontiers (Intensive and Extensive). In the Split-panel model we perform two different estimations: standard *OLS* and *OLS* with farm dummy variables. The inclusion of individual effects is not possible in the conventional Panel model because the probability of class membership is constant over time.

We find common patterns of results for both models and estimators. The probability of being in the intensive class increases the level of technical efficiency with respect to both frontiers for the two models and the two estimation methods. The coefficient of the time dummy indicates that the index of technical efficiency estimated using the extensive frontier increases over time while the index of technical efficiency estimated using the intensive frontier decreases over time. This result is probably due to different patterns across frontiers of the yearly shifts of the production frontier measured by the coefficients of the time dummy variables. The time dummy coefficients of the intensive frontier (Table A1 and A2 in the appendix) show a clear upward trend at the beginning of the period followed by a fall in the last few years. The time upward shift of the production frontier is compatible with decreasing technical efficiency if the movements of the frontier are due to productive improvements of a subset of leading farms while other farms do not move immediately towards the shifted frontier. On the other hand, the extensive frontier features smaller and erratic shifts over time. In our view, it is not surprising to observe farms approaching, on average, a production frontier with no sudden upper shifts.

Additionally, the Split-panel model provides a subtle result: the probability of being in the intensive class changes across farms and increases over time (on average). In other words, there are two sources of variation. The coefficient of the probability using plain *OLS* is estimated using both sources of variation. However, the coefficient of the probability using *OLS* with farm dummy variables is estimated using only the changes over time of probability. The results show that the estimates of the coefficient of probability are 0.0545 (intensive) and 0.0439 (extensive) when using only changes over time, while the same estimates increase to 0.0786 (intensive) and 0.0752 (extensive) when using both cross-section and time variation of the posterior probability. In summary, it seems that the bulk of the change of efficiency caused by changes in the probability of being in the intensive class can be attributed to changes over time of this probability.

4. Conclusions

The assumption of time-invariant prior probability of a latent class can be circumvented by estimating a single *LCM* but splitting the original panel into two or more periods. This proposal allows for the estimation of the technology of different production systems without an assumption that becomes increasingly untenable as the period of years analyzed increases. By doing so, we find in our empirical application that the probability of being in the "intensive dairy" class increases over time. This result can be interpreted as evidence of dairy farming intensification over the sample period.

We find differences in technical efficiency if we split the sample in terms of the production system using the posterior probability of each latent class. More precisely, the average technical efficiency is higher for farms that belong to the intensive latent class. Additionally, the intensive frontier dominates the extensive frontier, indicating that the intensive technology is more productive that the extensive one. This result can be seen as an economic rationale for the observed trend towards the intensification of dairy farms.

What are the policy implications of these findings? Given that the intensive technology is more productive and that intensive farms are more efficient, i.e. produce closer to their frontier than extensive farms, it seems that the trend towards intensification will continue in the near future.

The new reform of the Common Agricultural Policy may also affect farmers' technology choices. On the one hand, the phasing out of milk quotas may result in higher production. In this sense, intensive farms will find it easier to boost production since they do not depend heavily on forage. On the other hand, the new direct payment scheme which will move towards a uniform payment per hectare may lower the incentives to adopt intensive systems.

References

Aigner, D., K. Lovell, and P. Schmidt (1977), "Formulation and estimation of stochastic frontier production function models", *Journal of Econometrics*, 6:21-37.

Álvarez, A., J. del Corral, D. Solís and J.A. Pérez (2008), "Does Intensification Help to Improve the Economic Efficiency of Dairy Farms?", *Journal of Dairy Science*, 91(9):3693-3698.

Álvarez, A. and J. del Corral (2010), "Identifying Different Technologies Using a Latent Class Model. Extensive versus Intensive Dairy Farms", *European Review of Agricultural Economics*, 37(2):231-250.

Greene, W. (2005), "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model", *Journal of Econometrics*, 126:269–303.

Greene, W. (2007), *Limdep Version 9.0*. New York: Econometric Software.

Hallam, D., and F. Machado (1996), "Efficiency analysis with panel data: a study of Portuguese dairy farms", *Eur. Rev. Agr. Econ.* 23:79-93.

Hennessy, T., L. Shalloo, and P. Dillon (2005), "The economic implications of complying with a limit on organic nitrogen in a decoupled policy environment - an Irish case", *J. Farm Management*, 12:297-311.

Hoch, I. (1962), "Estimation of production function parameters combining time-series and cross-section data", *Econometrica*, 30:34-53.

Kompas, T. and T. Nhu Che (2006), "Technology choice and efficiency on Australian dairy farms", *Australian Journal of Agricultural and Resource Economics*, 50(1): 65–83.

Kumbhakar, S. and C. A. K. Lovell (2000), *Stochastic Frontier Analysis*, Cambridge: Cambridge University Press.

Newman, C. and A. Matthews (2006), "The productivity performance of Irish Dairy farms 1984-2000: A multiple output distance function approach", *J. Prod. Anal.*, 26:191-205.

Nehring, R., J. Sauer, J. Gillespie and C. Hallahan (2011), Intensive versus Extensive Dairy Production Systems: Dairy States in the Eastern and Midwestern U.S. and Key Pasture Countries the E.U.: Determining the Competitive Edge. Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Corpus Christi, TX.

Orea, L. and S. Kumbhakar (2004), “Efficiency measurement using a latent class stochastic frontier model”, *Empirical Economics,* 29:169–183.

Simpson, J. R., and J. H. Conrad (1993), “Intensification of cattle production systems in Central America: why and when”, *J. Dairy Science*, 76:1744-1752.

Tauer, L. and K. Belbase (1987), “Technical Efficiency of New York Dairy Farms”, *Northeastern Journal of Agricultural and Resource Economics,* 16:10-16.

Appendix

Table A.1: Estimation of the Panel Model

	Class 1		*Class 2*	
Variable	***Coefficient***	***Std. Error***	***Coefficient***	***Std. Error***
Constant	12.5724	.0158	12.542	.0152
Cows (lnx_1)	.7176	.0213	.4626	.0241
Feed (lnx_2)	.2969	.0137	.3623	.0152
Crop (lnx_3)	.0405	.0067	.0718	.0080
Animal (lnx_4)	.0296	.0091	.0964	.0110
Land (lnx_5)	.0368	.0104	.0339	.0125
$(lnx_1)^2$	.2030	.1445	-.0854	.1096
$(lnx_2)^2$	.0257	.0486	-.1737	.0612
$(lnx_3)^2$	.0575	.0152	.0523	.0149
$(lnx_4)^2$	-.0731	.0247	.0069	.0320
$(lnx_5)^2$	.2857	.0398	-.0624	.0542
lnx_1lnx_2	-.3277	.0641	.0663	.0732
lnx_1lnx_3	-.0999	.0395	.0103	.0301
lnx_1lnx_4	.3738	.0532	-.0394	.0561
lnx_1lnx_5	-.2055	.0584	-.0263	.0657
lnx_2lnx_3	.1055	.0238	-.0608	.0219
lnx_2lnx_4	-.0960	.0321	.0827	.0369
lnx_2lnx_5	.1288	.0405	-.0160	.0413
lnx_3lnx_4	-.0307	.0149	.0025	.0164
lnx_3lnx_5	-.0426	.0196	.0196	.0231
lnx_4lnx_5	-.0619	.0275	.0138	.0282
TD_{00}	.0206	.0129	.0072	.0172
TD_{01}	.0366	.0131	-.0046	.0173
TD_{02}	.0358	.0130	.0243	.0175
TD_{03}	.0313	.0131	-.0181	.0173
TD_{04}	.0519	.0131	.0038	.0171
TD_{05}	.0707	.0132	.0359	.0172
TD_{06}	.0952	.0132	.0366	.0173
TD_{07}	.0839	.0132	.0378	.0175
TD_{08}	.0529	.0134	-.0156	.0175
TD_{09}	.0559	.0134	-.0275	.0177
TD_{10}	.0661	.0139	.0032	.0172
Sigma	.0909	.0078	.1537	.0064
Lambda	1.0506	.3484	2.7329	.3866
Prior probability equation				
Constant	-11.906	5.2881		
ln(cows/land)	1.7919	.6078		
Ln(concentrate/cows)	1.3510	.6600		

Table A.2: Estimation of the Split-Panel Model

	Class 1		Class 2	
Variable	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Constant	12.5754	.0135	12.536	.0146
Cows (lnx_1)	.7491	.0203	.4553	.0211
Feed (lnx_2)	.2737	.0148	.3685	.0149
Crop (lnx_3)	.0508	.0068	.0767	.0078
Animal (lnx_4)	.0206	.0093	.0789	.0118
Land (lnx_5)	.0214	.0115	.0295	.0136
$(lnx_1)^2$	.2943	.1448	-.0405	.1092
$(lnx_2)^2$	.0121	.0460	-.0964	.0645
$(lnx_3)^2$	.0732	.0147	.0410	.0134
$(lnx_4)^2$	-.0530	.0263	-.0447	.0331
$(lnx_5)^2$	.2786	.0400	.0407	.0609
lnx_1lnx_2	-.3067	.0613	-.0549	.0693
lnx_1lnx_3	-.1158	.0375	-.0217	.0316
lnx_1lnx_4	.2953	.0530	.0704	.0591
lnx_1lnx_5	-.2457	.0563	.0135	.0596
lnx_2lnx_3	.0825	.0226	-.0131	.0217
lnx_2lnx_4	-.0526	.0331	.0418	.0396
lnx_2lnx_5	.1130	.0410	-.0562	.0452
lnx_3lnx_4	-.0275	.0151	.0003	.0160
lnx_3lnx_5	-.0263	.0199	-.0053	.0224
lnx_4lnx_5	-.0350	.0274	.0306	.0289
TD_{00}	.0201	.0126	-.0013	.0162
TD_{01}	.0376	.0129	-.0108	.0174
TD_{02}	.0351	.0133	.0238	.0188
TD_{03}	.0283	.0130	-.0151	.0166
TD_{04}	.0471	.0132	.0099	.0165
TD_{05}	.0660	.0128	.0408	.0167
TD_{06}	.0917	.0130	.0457	.0167
TD_{07}	.0790	.0129	.0463	.0172
TD_{08}	.0531	.0131	-.0143	.0171
TD_{09}	.0563	.0131	-.0292	.0172
TD_{10}	.0697	.0134	.0007	.0170
Sigma	.0883	.0067	.1497	.0060
Lambda	1.1750	.3155	3.0344	.4630
Prior probability equation				
Constant	-13.112	4.2444		
ln(cows/land)	.8632	.3943		
Ln(concentrate/cows)	1.5513	.5247		