Soumission aux 10èmes Journées de Recherches en Sciences Sociales (JRSS)

Paris-La Défense – 8 et 9 décembre 2016

Heterogeneity in spatial interaction effects on farm survival and growth: evidence from Brittany

Legrand D. F. Saint-Cyr[*1], Hugo Storm[2], Thomas Heckelei[2], Laurent Piet[1]

*Corresponding author: SMART, AGROCAMPUS OUEST, INRA, 4 allée Adolphe Bobierre, CS 61103, F-35011 Rennes cedex; legrand.saintcyr@agrocampus-ouest.fr

[1]SMART, AGROCAMPUS OUEST, INRA, 35000, Rennes, France

[2]University of Bonn, Germany

Abstract

Accounting for spatial interdependence is relevant for the analysis of structural change in farming and for the assessment of policy changes because of potential interactions between farms. However, impacts of neighbouring farm characteristics may vary according to farmers' motivations. To identify specific farms' relationships considering farm survival and growth process, a mixture modelling framework that enables capturing heterogeneity in spatial interdependence between farms is developed. An application to a panel of farms in Brittany in France from 2004 to 2014 shows that relationships between farms are more in terms of competition for land than positive spill overs of new technology adoption, leading to a negative impact of neighbouring farms' size on the probability to survive for a majority of farms. The results also show that neighbouring farms' size has a positive but indirect effect on farm growth through its impact on farm survival. These results suggest that more attention should be paid to heterogeneity in spatial interdependence between farms for a better understanding of farm size dynamics.

Keywords: EM algorithm, Farm interdependence, Mixture model, Spatial interactions, Unobserved heterogeneity

JEL classification: C23, D22, Q12

1. Introduction

The farming sector faced considerable structural change over the last decades. In most developed countries, the total number of farms decreased significantly and their average size increased, implying changes in the distribution of farm sizes (Bollman *et al.*, 1995; Eastwood *et al.*, 2010). Understanding the factors that affect farmers' decisions to enter or exit farming as well as farm growth has been a concern of agricultural economists and policy makers for quite some time. For example, Shapiro *et al.* (1987) examined the relationship between farm size and growth in Canada over the period 1966-81; Sumner and Leiby (1987) analysed the effects of human capital on size and growth to a sample of Southern dairy farms in the United States; Weiss (1999) investigated farm survival and growth in Upper Austria; Breustedt and Glauben (2007) examined the exit process of Western European farmers; Dong *et al.* (2010) studied the exit decision of finisher hog producers in North America. In France, Aubert and Perrier-Cornet (2009) studied factors that influence survival and growth of small farms while Bakucs *et al.* (2013) investigated the relationship between size and farm growth. Among others, these studies identified important aspects of structural change in farming and showed that farm survival and growth processes may help understand farm dynamics in different farming contexts. More recently, Storm *et al.* (2015) empirically investigated the effects of direct payments on exit rates of Norwegian farms and showed that the spatial interdependence between farms is an important factor in farmers' decisions to maintain their production activities. The authors showed that accounting for spatial interdependence between farms may be highly relevant for an aggregate assessment of policy changes in agriculture.

The study presented in the present paper adds to the existing literature especially in three ways. Firstly, we account for spatial interdependence between farms in both farmers' decision to continue their production activities over time and the process of farm growth. Indeed, neighbouring farms' characteristics may influence farm survival as well as farm growth since these two processes may be related to each other when they impact farms which are close to each other. Accounting for interdependence in both farmers' decision to stay in business and farm growth process may therefore improve the analysis of structural change in farming. Some studies have already investigated both farm growth and survival in different farming contexts (Weiss, 1999; Aubert and Perrier-Cornet, 2009). However, none of the existing studies have taken into account the potential spatial interdependence between farms, to the best of our knowledge.

Secondly, we extend the existing methods by using a mixture modelling approach to investigate spatial interdependence between farms. Generally, studies in this strand of the literature estimate mean effects of neighbouring farms' characteristics on farmers' decision to exit farming or to increase their operated farm size (see Storm *et al.* (2015) for a recent example). The results from these studies are therefore based on the assumption that all farms will behave alike given the investigated characteristics. However, some farms may be more or less sensitive to the characteristics of their neighbours (e.g., direct payment received, farm size, etc.), due to some specific individual characteristics. If all of these characteristics were observed, controlling for them would lead to a more efficient estimation of the impacts of neighbouring farms' characteristics. Otherwise, the resulting parameters may be biased and inconsistent due to unobserved farm heterogeneity (Kyriazidou, 1997; Pennings and Garcia, 2004). One way to tackle this issue is to use modelling frameworks that allow controlling for unobserved farm heterogeneity. Various modelling approaches such as fixed and random effect, random parameter and mixture models can be used to control for unobserved heterogeneity (Greene, 2012). Holloway *et al.* (2007) reported several strategies that could be adopted in order to control for unobserved heterogeneity in modelling spatial dependence. Among these strategies, these authors argued that the mixture modelling framework seems to be the most attractive since it is simple and intuitive. According to them, one of the main advantages of a mixture modelling approach is that it allows the data themselves to sample select and it gathers observations characterised by similar relations between dependent and independent variables. The mixture modelling approach can group farms with similar behaviours and therefore could help identify specific impacts of neighbouring farm's characteristics.

Thirdly, we develop the mixture modelling approach in order to handle panel data to capture potential dynamic effects in farmers' decisions. Some studies in the literature have already used panel data to study farmers' decisions to exit farming and farm growth process (see Bakucs *et al.* (2013) for a recent example). However, to the best of our knowledge, we are the first to investigate spatial farm interdependence both using panel data and controlling for unobserved heterogeneity. The fundamental advantage of a panel dataset over a cross section one is that the former allows greater flexibility in modelling differences in the behaviour of individuals (Greene, 2012). We can therefore expect that using a mixture modelling approach could group farms with similar behaviours and thus reveal different impacts of neighbouring farms' characteristics on farmers' decision to exit farming or to increase their production capacity.

This paper is structured as follows. The next section provides theoretical arguments supporting the empirical application of this study. Sections 3 and 4 present the modelling approach and the corresponding estimation procedure, respectively. The data used for our empirical application and explanatory variables for the model specification are presented in Section 5. Section 6 reports the main results. The last section concludes with some considerations on possible improvements of this study for further research.

2. On heterogeneity and spatial interdependence between farms

Neighbouring farms' characteristics may have important impacts on own farm size and/or on farmers' decision to exit farming. According to Storm *et al.* (2015), a farm will survive if its willingness to pay (WTP) for land is greater than the WTP for land of its neighbours. As the WTP for land of a farm depends on the farm characteristics, the farmer's decision resulting from the difference in his/her WTP for land is therefore related to his/her neighbouring farms' characteristics. In this study, we argue that the impact of the characteristics of neighbouring farms also depends on the own characteristics of the farmer under consideration. Focusing on the neighbouring farms' specific characteristic that is size, we extend Storm *et al.* (2015)'s theoretical background providing some additional elements supporting this proposition.

The existing literature distinguishes two types of effects of neighbouring farms' size originating from technology adoption. On the one hand, neighbours can be viewed as competitors especially for the acquisition of plots (Weiss, 1999). In this case, a farmer surrounded by larger farms may be constrained to close his/her operation since larger farms are more likely to adopt new technologies earlier given their potentially greater access to information and better financial capacity (Goddard *et al.*, 1993). Larger neighbours therefore have a higher WTP for land, leading to a negative impact on the probability to survive and to enlarge, for the farm under consideration. On the other hand, neighbours can be considered as a source of motivation and example to adopt new technologies (Case, 1992; Holloway *et al.*, 2002). In this case, neighbouring farms' size positively influences the survival of the farm under consideration, because a farmer surrounded by larger farms is more likely to benefit from the innovation of larger neighbouring farms (Harrington and Reinsel, 1995). This may imply an increase in the WTP for land for those neighbouring farms since new technology adoption generally requires acquisition of land for an optimal use of the technology.

However, these interactions among neighbours may depend on the farm under consideration. Indeed, we expect that the effect of neighbouring farms' size is rather heterogeneous across farms under consideration, and crucially depends on the type and characteristics of the farm and the farmer considered. One of the most important sources of farm heterogeneity that may shape farmers' behaviour is their motivation: neighbouring farms' size is more likely to have an impact (positive or negative) on farmers who are mainly motivated by profit maximisation. In the context of a free market competition, such business-oriented farms are constrained either to innovate or to exit, leaving resources to be acquired by more innovative competitors in the latter case (Harrington and Reinsel, 1995; Jackson-Smith, 1999). The persistence of commercial farms thus depends on their competitiveness, that is, on their capacity to innovate. However, this capacity differs across farms and depends on a variety of factors such as accessibility to technology and land, managerial capacity, risk perception, attitudes towards risk, etc. (Bowman and Zilberman, 2013; Conradt *et al.*, 2014; Trujillo-Barrera *et al.*, 2016).

Nevertheless, not all farmers give priority to the commercial aspect of farming activities (Maybery *et al.*, 2005; Howley *et al.*, 2014). Some non-competitive farmers may keep their activity because of prevailing non-pecuniary motives (Harrington and Reinsel, 1995), perhaps enjoying the farming lifestyle (Hallam, 1991), and may maintain production at sub-optimal levels (Howley, 2015). For example, it may be the case for some environmentally oriented farms (Willock *et al.*, 1999) or certain hobby farms (Daniels, 1986; Holloway, 2002). For such kinds of farms, the new technology is evaluated on aspects other than financial viability (Mzoughi, 2011). Then, technology is adopted only if it is considered as conform to some predefined criteria that are set by the farmer based on his/her non-pecuniary objectives. Overall, one can thus expect a lower or even no specific impact of neighbouring farms' size on the probability to survive or to increase the operated size, for those farmers characterised by prevailing non-pecuniary motives.

Based on these considerations, we hypothesise that there are at least two different types of farms that respond differently to neighbouring farms' size: a negative response because of competition for land, or a positive response resulting from positive spill overs of new technology adoption. Two questions arise from this: which farms are more likely to be in each of these specific types? And which type of relationship is the prevailing one in a specific farming context? Investigating both questions may help understand farm size dynamics in specific farming contexts. In the following, we apply a modelling framework that enables identifying specific types of relationships between farms and thus contributes to answer the second question. Even though, we identify the main characteristics of farms in each of these types in a second step, the first question remain out of the scope of this study and would need to be investigated in a more efficient way in future works.

3. Modelling approach

We separately investigate the impacts of neighbouring farms' size on farm survival on the one hand, and on the process of farm growth on the other hand.

Regarding farm survival, a probit model is applied. A latent regression underlies the probit model, where the latent variable represents the utility that is obtained from staying in or exiting the farming sector. Farmers' utility may be affected by their own WTP for land as well as their neighbours' WTP for land. The latent variable y^*_{it} underlying the probit model determines the outcome of the farmer's decision to stay in business in two consecutive years. As yearly information about farmers' decisions is available, the observed outcome can be thus obtained as:

$$\begin{aligned} y_{it} = 1 \quad &\text{if} \quad y^*_{it} > 0, \quad \forall t \in T_i \\ y_{it} = 0 \quad &\text{if} \quad y^*_{it} \leq 0 \end{aligned} \tag{1}$$

where y_{it} is the observed outcome at time t for farm i which takes values: $y_{it} = 1$ if the farm survives two consecutive years, and $y_{it} = 0$ otherwise; T_i is the length of time that farm i is observed. The latent variable at time t is in turn given by:

$$y^*_{it} = \mathbf{x}_{it-1}\boldsymbol{\beta} + \epsilon_{it}, \qquad t = 1, 2, \cdots, T_i \leq T \tag{2}$$

where $\boldsymbol{\beta}$ are the parameters to be estimated, $\mathbf{x}_{(it-1)}$ are own and neighbouring farm characteristics. The disturbances ϵ_{it} are T-variate, normally distributed with $T \times T$ positive definite covariance matrix $\boldsymbol{\Sigma}$. The typical element of $\boldsymbol{\Sigma}$ is denoted σ_{ts} and the standard deviations σ_t. The data on $\mathbf{x}_{(it-1)}$ are assumed to be strictly exogenous, which implies that $\mathrm{Cov}[\mathbf{x}_{it-1}; \epsilon_{js}] = 0$ across all individuals i and j and all periods t and s (see Greene (2004) for more details).

Farm growth is represented thanks to a simple linear regression model. The total land used at any specific time t is thus given by:

$$y_{it} = \mathbf{x}_{it-1}\boldsymbol{\theta} + u_{it}, \qquad t = 1, 2, \cdots, T_i \leq T \tag{3}$$

where θ is the vector of parameters to be estimated, $x_{(it-1)}$ are own and neighbouring farms' characteristics, and u_{it} is an *iid* normally distributed error term.

In both the probit and the simple linear regression models, the explanatory variables are lagged one year to reflect the response delay of the adjustment to exogenous variables. Neighbouring farms' characteristics are introduced in the specification of the models to capture spatial effects and interdependence between farms. It should be noted that spatial interdependence between farms are captured using explanatory variables defined at certain geographical level, instead of defining spatial weighting matrix which is the methodology generally applied in the literature (see Section 5.1). This approach is convenient for our estimation procedure and has already been applied to account for spatial dependence in other strands of the economic literature (see Teillard *et al.* (2012) and Allaire *et al.* (2015) for recent examples).

As argued in the previous section, neighbouring farms' size may influence farmers' decisions in various ways. To capture the heterogeneity in farmers' responses to their neighbouring farms' characteristics, we apply a mixture modelling approach, which allows capturing unobserved heterogeneity. The mixture modelling approach supposes that the farm population is divided into more than one homogeneous group; each type of farms is characterised by a specific effect of the exogenous variables, including neighbouring farms' size, on farmers' decisions. Let $\mathbf{y} = (\mathbf{y}_1^T, \cdots, \mathbf{y}_n^T)$ denote the observed random sample where $\mathbf{y}_i$ is the sequence of choices or states of farm i over a certain period of time. Under a mixture approach, the density $f(\mathbf{y}_i)$ is written as (McLachlan and Peel, 2004):

$$f(\mathbf{y}_i) = \sum_{g=1}^{G} \pi_g f_g(\mathbf{y}_i) \tag{4}$$

where π_g is the proportion of farms belonging to type g with $g = 1, 2, \cdots, G$, and f_g is type-g density as described by equations (2) and (3) for farm survival and growth process, respectively. Since the unobserved types have to be exhaustive and mutually excluding, the π_g proportions are non-negative and sum up to unity.

Under such a mixture approach, the conditional probability density for the observed data for farm i is:

$$f(\mathbf{y}_i|\mathbf{X}_i; \mathbf{\Psi}) = \sum_{g=1}^{G} \pi_g f_g(\mathbf{y}_i|\mathbf{X}_i; \mathbf{\Phi}_g) \tag{5}$$

where $\mathbf{\Psi} = (\pi_1, ..., \pi_G, \mathbf{\Phi}_1, ..., \mathbf{\Phi}_G)$ are the parameters to be estimated with $\mathbf{\Phi}_g = \boldsymbol{\beta}_g$ for the probit model and $\mathbf{\Phi}_g = \boldsymbol{\theta}_g$ simple linear regression model; and f_g is the respective probability density function specific to farm type g, given by:

$$f_g(\mathbf{y}_i|\mathbf{X}_i; \mathbf{\Phi}_g) = f(\mathbf{x}_{it-1}; \boldsymbol{\beta}_g) = \left[F(\mathbf{x}_{it-1}\boldsymbol{\beta}_g)\right]^{y_{it}} \left[1 - F(\mathbf{x}_{it-1}\boldsymbol{\beta}_g)\right]^{(1-y_{it})} \tag{6}$$

for the probit model where $F(\mathbf{x}_{(it-1)}; \Phi_g)$ is the cumulative density function for and farm type g and y_{it} is the observed outcome. For the simple linear regression model, the probability density function writes:

$$f_g(\mathbf{y}_i|\mathbf{X}_i; \mathbf{\Phi}_g) = f(\mathbf{x}_{it-1}; \boldsymbol{\theta}_g) = \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp\left\{-\frac{1}{2\sigma_g^2}(y_{it} - \mathbf{x}_{it-1}\boldsymbol{\theta}_g)^2\right\} \tag{7}$$

where σ_g^2 is the variance of the error term specific to farm type g, and y_{it} is the dependent variable.

4. Estimation of the models

The mixture models described in the previous section are estimated using the maximum likelihood method. Assuming that, for each model, observations are independent within farm types given $\mathbf{x}_{(it-1)}$, the log-likelihood (LL) function for the parameters $\mathbf{\Psi}$ of the model, conditional on observing $\mathbf{y}_i$, is written as:

$$LL(\mathbf{\Psi}) = \sum_{i=1}^{N} \ln \left\{ \sum_{g=1}^{G} \pi_g \prod_{t=1}^{T_i} f(\mathbf{x}_{it-1}; \mathbf{\Phi}_g) \right\} \tag{8}$$

4.1. Implementing the Expectation-Maximisation (EM) algorithm

As the type of farms is unknown beforehand, the expectation-maximisation (EM) algorithm is used to estimate the parameters of the models. The EM algorithm simplifies the complex log-likelihood in equation (8) into a set easily solvable log-likelihood functions by treating the unobserved farm type as a missing information (McLachlan and Krishnan, 2007). Using a non-parametric mixing distribution approach, the complete log-likelihood is thus written as (Train, 2008):

$$LL_c(\mathbf{\Psi}) = \sum_{i=1}^{N}\sum_{g=1}^{G} v_{gi}\ln\left\{\pi_g \prod_{t=1}^{T_i} f(\mathbf{x}_{it-1};\mathbf{\Phi}_g)\right\} \tag{9}$$

In this case, v_{ig} is called the 'posterior' probability that farm i belongs to type g given y_i, that is $P(v_{gi} = 1|\mathbf{y}_i)$, while π_g is the 'prior' probability of the mixture (McLachlan and Peel, 2004). The expression in equation (9) can then be divided into two components:

$$\begin{aligned} LL_1 &= \sum_{i=1}^{N}\sum_{g=1}^{G} v_{gi}\ln\pi_g \\ LL_2 &= \sum_{i=1}^{N}\sum_{g=1}^{G} v_{gi}\sum_{t=1}^{T_i}\ln f(\mathbf{x}_{it-1};\mathbf{\Phi}_g) \end{aligned} \tag{10}$$

As the farm type is not observed, the posterior probability that farm i belongs to type g (*i.e.*, v_{ig}) has to be estimated from the observations. The EM algorithm therefore consists of the four following steps:

(i) Initialisation: Arbitrarily choose initial values $\mathbf{\Psi}^0 = (\pi_1^0, ..., \pi_G^0, \mathbf{\Phi}_1^0, ..., \mathbf{\Phi}_1^0)$ for the parameters of the models.

(ii) Expectation: At iteration $p+1$ of the algorithm, compute the expected probability that farm i belongs to a specific type g while observing $\mathbf{y}_i$ and given the parameters $\mathbf{\Psi}^{(p)}$. This conditional expectation probability, that is, the posterior probability $v_{gi}^{(p+1)} = v_{gi}(\mathbf{y}_i;\mathbf{\Psi}^p)$, can be obtained according to the Bayes' law:

$$v_{gi}^{(p+1)} = \frac{\pi_g^{(p)}\prod_{t=1}^{T_i} f(\mathbf{x}_{it-1};\mathbf{\Phi}_g^{(p)})}{\sum_{h=1}^{G}\pi_h^{(p)}\prod_{t=1}^{T_i} f(\mathbf{x}_{it-1};\mathbf{\Phi}_h^{(p)})} \tag{11}$$

Replacing v_{ig} by its expected value in equation (9) leads to the conditional expectation of the complete data log-likelihood.

(iii) Maximisation: Update $\mathbf{\Phi}^p$ by maximising the complete log-likelihood conditional on the observations. The model parameters are thus updated as:

$$\mathbf{\Phi}^{(p+1)} = argmax_{\mathbf{\Phi}}\sum_{i=1}^{N}\sum_{g=1}^{G} v_{gi}^{(p+1)}\sum_{t=1}^{T_i}\ln f(\mathbf{x}_{it-1};\mathbf{\Phi}_g) \tag{12}$$

The maximisation process of equation (12) is straightforward. The parameters of the model ($\hat{\mathbf{\Phi}}^{(p)}$) are updated considering $v_{gi}(\mathbf{y}_i;\mathbf{\Psi}^{(p)})$ as weighting factors for each observation. Then, the prior probability of belonging to type g is updated as:

$$\pi_g^{(p+1)} = \frac{\sum_{i=1}^{N} v_{gi}^{(p+1)}}{\sum_{i=1}^{N}\sum_{h=1}^{G} v_{hi}^{(p+1)}}, \qquad \forall g \in G \tag{13}$$

(iv) Iteration: Return to expectation step (ii) using $\boldsymbol{\pi}^{(p+1)}$ and $\mathbf{\Phi}^{(p+1)}$ and iterate steps (ii) and (iii) until convergence of the observed log-likelihood given by equation (8).

At convergence, the resulting parameters ($\hat{\mathbf{\Psi}}$) are considered as optimal. Because of the potential presence of a high number of local maxima (Hess *et al.*, 2006), the EM algorithm is run several times with various randomly chosen initial values, and those providing the largest likelihood at convergence are chosen as the best ones.

4.2. Choosing optimal number of farm types

Despite the intuition about relationships between farms as described in Section 2, we have no a priori information about the optimal number of homogeneous farm types that may exist in a specific farming context. The total number of components for the mixture of probit models as well as for the mixture of linear regression models are thus chosen based on information criteria. The selected criteria are derived on the resulting value of the log-likelihood of the corresponding model $LL_G(\mathbf{y}; \hat{\mathbf{\Psi}})$ for a total of G homogeneous types. The basic principle under these information criteria is parsimony, that is, all other things being the same, the model with fewer parameters is preferred (Andrews and Currim, 2003). The selection criteria thus derive from the following formula:

$$C_G = -2\left\{LL_G(\mathbf{y}; \hat{\mathbf{\Psi}})\right\} + \kappa N_G \tag{14}$$

where $LL_G(\mathbf{y}; \hat{\mathbf{\Psi}})$ is the overall population log-likelihood value computed with the resulting estimated parameters for the model specified with G types, N_G is the total number of free parameters in the model and κ is a penalty constant.

Depending on the value chosen for κ, we obtain the well-known Akaike Information Criterion (AIC) if κ=2, and Bayesian Information Criterion (BIC) if $\kappa = \log(N)$ where N is the total number of observations. However, based on Monte Carlo simulations, it has been proved that the modified AIC (AIC3) with κ=3 is preferable to the basic AIC and BIC since the two latter criteria more severely penalise the addition of parameters than AIC3 (Andrews and Currim, 2003). For all of these heuristic criteria, smaller values mean more parsimonious models.

5. Empirical application

5.1. Data used

For our empirical application, we used data provided by the 'Mutualité Sociale Agricole' (MSA), the French authority for farmers' healthcare and social security. The MSA database contains information about all individuals who declare carrying out a non-salaried farming activity in France, and about their farm. Information is collected annually and is available for farmers who were active on January 1st of each year, from 2004 to 2014. The database can be actually considered as almost exhaustive for the French farm population∗, so we can assume that a farm: i) survived if it remained in the MSA database over the whole period of observation; ii) started business if it entered the database after 2004; iii) quit farming if it was not in the database before 2014.

Using the MSA data requires several preliminary treatments to prepare the database before analysis. Here we only mention the most significant ones. First, the data have to be consolidated at the farm level because the MSA collects information at the affiliated physical person level, *i.e.*, the farm holder level and not the farm itself. This is made possible because, in the database, each individual person is assigned the farm number he/she belongs to. Second, the utilised agricultural area (UAA) has to be aggregated with care at the farm level because it is not simply the sum of the areas reported for each partner of the farm. Indeed, the recorded UAA at the individual partner level is calculated with respect to the proportion of the total social shares of the farm he/she holds. Then, whenever one or several partners are not affiliated to the MSA because they are external to the farming sector, social shares in the database do not sum up to unity, and hence the total UAA has to be computed taking into account that a part of it accrues to partners who are not observed in the database. Third, assumptions have to be made for other variables when consolidated at the farm level in the case where the farm is run by several partners. This is typically the case for some of the variables used in the analysis: age, farming specialisation and legal status. As for age, a choice has to be made on the 'age' of which person characterises the farm. Here, we arbitrarily chose to retain the median age of the farm's partners. As for the farm production specialisation, it is determined from the category of professional risks each partner is registered to. In the database, there exists 16 such risk categories such as 'cereals and industrial crops', 'dairy cattle breeding', 'pig farming', 'wine growing', etc. Then, when the farm includes several production units (*e.g.*, a 'crop' unit plus a 'livestock' unit in dairy farms), each partner may subscribe for only one of the corresponding professional risks depending on the unit he/she is specialised in. Therefore, several such risk categories may coexist on the same farm and an assumption has to be made on how to aggregate them to avoid classifying such farms as 'mixed farms'. Here, we chose to assign to the farm the risk category which represented the two thirds of partners or, whenever such majority did not exist, to classify the farm as mixed. In a last step, the MSA risk categories were translated into 13 'types of farming' (farm production specialisation) chosen in the nomenclature used by the French agricultural statistics office for the agricultural census and related surveys. Finally, as regards the legal status, while it should be recorded identically for the different partners of a specific farm, it appeared that this was not so in some cases. In such situations, we assigned

∗The database is considered as 'almost' exhaustive because only it does not survey small farms which do not contribute to the MSA as well as corporate farms employing only salaried workforce.

the predominant legal status (*i.e.*, the mode of the observed values) or, when this was not possible because of several modes, we assigned the status corresponding to the higher degree of incorporated form of association.

In this empirical application, we restricted our investigation to farms located in Brittany (Western France), which is one of the largest agricultural regions in France.

5.2. Dependent and explanatory variables

For the analysis of farm survival, the dependent variable takes the value 1 if the farm survives and 0 otherwise. Since we consider all farms whatever their production specialisation or legal status, a farm is said to survive from one year to the next if it remains present in the database these two consecutive years. For the analysis of farm growth, the dependent variable consists in the change in the total UAA in hectares over the period of observation, with an increase representing a positive growth or enlargement, and a decrease representing a negative growth or decline.

The analysis of the spatial interdependence between neighbouring farms in the process of farm survival and growth requires special attention because the MSA database exhibits two main limitations for such a study.

Firstly, the MSA database contains only a few variables that can be used to explain farm survival and growth. We thus choose to concentrate on the possible impacts of the limited set of available variables. Other databases are merged with the MSA to provide additional information especially at different special scales. The most important farm characteristic that may play a role in the probability to survive is farm size in terms of total UAA (*area*) and farm total agricultural profit (*agri_profit*). Both explanatory variables are expected to positively influence the probability to survive and to increase the operated farm size because such farm characteristics may increase the farm's WTP for land. While the total land used is rather a proxy of path dependency, the total agricultural profit indicates whether or not farming is a profitable activity. The age of the farm holder (*age*), dummies indicating that the farm production specialisation is pig and/or poultry (*pig/poultry*), and a dummy indicating that the legal status of the farm is a corporate farm in opposition to partnerships or individual farms (*corporate*), are also included in the model specification. Theses variables are introduced to capture farm observed heterogeneity. Age square is used to capture non-linear effect of the farm holder age. Indeed, the age of farmers may be positively related to the probability of surviving and of increasing the operated size, since farmers' experience and skills may increase over years; by contrast, older farmers especially close to retirement time, may be less motivated to either compete for land or to adjust their operated size over time, leading to the opposite effect. As farm specialisation in pig and/or poultry generally requires less land, this type of farms may be less likely to compete for new plots. We thus expect a positive effect on farm survival and a negative impact on farm size change in term of total land used. Corporate farm legal status is also supposed to be positively related to farm survival and growth because these farms are generally in a better position to compete for land since they may have lower financial and credit constraints. As a farm's WTP for land may decrease at retirement time despite high agricultural profits, we control for the impact of retirement time by using an interaction term between farm agricultural profit and a dummy indicating that the farmer is close to retirement time (*agr_profit_×_retirement*). According to the MSA, the minimum age for retirement in France is 60 years old but farmers' behaviour may change earlier. Since some studies have indeed shown that farmers' succession is prepared between 5 and 10 years in advance, we choose to retain 55 years old and above as the indicator of retirement closeness (Gaté and Latruffe, 2016).

Secondly, the MSA database contains no information about the precise geographical location of the farmstead and farm plots. It is therefore impossible to determine the actual distance between farms. Only the municipality where the farmstead is located is available in the database. As municipalities in France are relatively small and given the dispersion of farm plots on French farms, farms may compete for land in their own municipality and even in neighbouring municipalities (Piet and Cariou, 2014; Latruffe and Piet, 2014). We thus use average farm characteristics at the municipality level to capture the effects of neighbouring farms' size on a farm's survival and growth. At a first spatial scale, we consider farms located in the same municipality as the farm under consideration. Brittany counts 1,270 municipalities with an average area of 21 square km. From this, we calculate the average farm size by municipality *(average_mun_area)* and use it as a proxy for neighbouring farms' size. We also calculate, using the MSA database, the average age of farm holders (*average_mun_age*), the share of farms specialised in pig and/or poultry (*mun_pig/poultry_share*) and the share of corporate farms (*mun_corporate_share*) at the municipality level.

Following Storm and Heckelei (2016), we also include the same variables calculated at a larger spatial scale than the municipality. This allows distinguishing the effects of farm interactions that take place on a smaller spatial scale from spatial correlation arising from unobserved spatially correlated regional characteristics at a larger scale. Specifically, we calculate the average characteristics and shares for small agricultural regions (SAR), which is a geographical unit that may contain one or more municipalities. The SAR level is a zoning that was specifically designed to define units with homogeneous conditions in terms of agricultural systems, soil and climate. The mean size of a French SAR is 22.4 ±13 square km (Teillard *et al.*, 2012). Based on the INSEE 2007 classification, there exist 25 SAR in Brittany, that is, about 50 municipalities by SAR on average. The variables (farm area, age of farm holder, pig/poultry specialisation and

Table 1. Definition and descriptive statistics of explanatory variables (n=315,529)

Variable	Code	Mean	St.Dev.	Min	Max
Farm level					
Age of the farm holder (years)	*age*	48.45	9.12	18.50	99.00
Total UAA (ha)	*area*	48.82	41.20	0.00	580.30
Total agricultural profit (1,000 Euros)	*agri_profit*	10.78	12.72	-313.92	465.72
Pig/poultry specialisation dummy (1 if yes)	*pig/poultry*	0.18	0.38	0.00	1.00
Corporate farm dummy (1 if yes)	*corporate*	0.46	0.49	0.00	1.00
Municipality level (mun)					
Average farm holder age	*average_mun_age*	48.45	2.33	25.00	88.00
Average farm size	*average_mun_area*	48.82	13.60	0.00	227.29
Share of pig/poultry farms (%)	*mun_pig/poultry_share*	18.00	13.00	0.00	100.00
Share of corporate farms (%)	*mun_corporate_share*	46.00	14.00	0.00	100.
Small agricultural region level (sar)					
Average farm holder age	*average_sar_age*	48.45	1.12	44.30	51.28
Average farm size	*average_sar_area*	48.82	7.66	13.92	70.61
Share of pig/poultry farms	*sar_pig/poultry_share*	18.00	7.00	1.00	29.00
Share of corporate farms	*sar_corporate_share*	46.00	8.00	24.00	70.00
Employment regional level					
Unemployment rate (%)	*unempl_rate*	7.03	1.21	3.70	9.90

Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

corporate legal status) are defined here at the SAR level as: *average_sar_area*, *average_sar_age*, *sar_pig/poultry_share* and *sar_corporate_share*.†

Additionally, we use the rate of unemployment in employment regions (*unempl_rate*). The unemployment rate captures the opportunities for off-farm activities and is thus supposed to have a direct effect on the probability for farms to remain in farming, and only an indirect effect on farm growth through its impact on farm survival. A time trend is used in addition. It may capture potential effects of, for example, technical change in farming that may influence farm survival and growth.

We use the same set of explanatory variables to explain both farm survival and growth. In the second model describing farm growth, the logarithm of the total UAA is used instead of total UAA *per se* to facilitate convergence of the model. All explanatory variables are lagged one year because it is supposed that farmers take their decisions based on available information in the year just preceding. Descriptive statistics are reported in Table 1.

Before starting with the results, it should be finally noted that the mixture of probit and the mixture of linear regression models are estimated separately. As explanatory variables are

†It should be noted that farm agricultural profit is not defined here neither at the municipality level nor at the SAR level since it is found highly non-significant at both spatial levels.

lagged one year for both models, the estimations are thus performed using 317,177 and 294,288 observations out of the 344,617 in the database, respectively for the models of farm survival and farm growth. In the following section, the results are thus presented for each model separately.

6. Results

6.1. Farm survival model

Table 2 reports the estimated parameters for both a pooled estimation where unobserved heterogeneity is not considered and the mixture probit model. Estimated parameters of the pooled probit thus constitute a mean effect of the considered farms' own characteristics and neighbouring farms' characteristics on the probability to survive from one year to the next, while the mixture model identifies impacts which are specific to the endogenously determined homogeneous farm types.

The results from the homogeneous model are consistent with our expectations. Overall, a positive impact is observed for the age of farm holders, the operated farm size (land), and the total agricultural profit. The results show a non-linear impact both for the age of farm holders and the total farm area. The negative impact of the square of age means that older farm holders are less likely to remain active over years. The effect the square of farm area is lower which may suggest that very large may face some constraints that tend to decrease their probability to survive in comparison to smaller farms. A positive effect is also observed for farm specialisation in pig and/or poultry and for farms operated under a corporate legal status. This result is in agreement with our expectations: the probability to survive of farm specialised in pig/and or poultry production may be less related to competition for land, while corporate farms may be in a better place to compete because of lower financial and credit constraints. Farm agricultural profit is also found to positively affect farm survival, but farm holders close to retirement time tend to leave the farming sector although this activity is profitable may be because they expect to receive good pension at this time. The average farm size at the municipality level is not significant which may suggest that ignoring farm heterogeneity is not appropriate to analyse the impact of neighbouring size. However, the probability to survive is positively related to the average farm size at the small agricultural region level, which indicates unobserved spatial correlation between regional characteristics.

The mixture probit model distinguishes three optimal types in the studied farm population, especially differing with respect to the effect of neighbouring farms' size. Across all farms, the effect of neighbouring farms' size is negative but insignificant. However, the first and the second types of farms are characterised by a significant positive and, respectively, negative impact of neighbouring farms' size on the probability to survive. In the third type, the effect is considerably smaller and not significant. The negative influence of neighbouring farms' size on the probability

Table 2. Estimated parameters for both the pooled and the mixture probit model for farm survival

Variable code	Pooled	Mixture		
		type 1	type 2	type 3
intercept	0.0358	1.3314**	-0.8903*	-60.2809***
	(0.3396)	(0.4774)	(0.3875	(1.3082)
time_trend	0.0062**	-0.0222***	-0.0286***	0.1037***
	(0.0028)	(0.0040)	(0.0033)	(0.0092)
age	0.0104***	-0.0258***	0.0066*	3.2279***
	(0.0025)	(0.0036)	(0.0027)	(0.0328)
age_square	-0.0003***	2.10e-05	-0.0001***	-0.040***
	(2.23e-05)	(3.26-05)	(2.51e-06)	(0.0004)
area	0.0042***	-0.0173***	0.0092***	0.0007
	(0.0002)	(0.0004)	(0.0002)	(0.0008)
area_square	9.64e-06***	0.0001***	2.40e-05***	4.37e-05***
	(1.09e-06)	(2.29e-6)	(1.30e-06)	(6.01e-06)
agri_profit	0.0009**	-0.0499***	0.004***	0.0508***
	(0.0004)	(0.0007)	(0.0004)	(0.0045)
agri_profit_×_retirement	-0.0185***	-0.0283***	-0.0162***	-0.0515***
	(0.0006)	(0.0008)	(0.0006)	(0.0046)
pig/poultry	0.0228**	0.2074***	0.0281*	0.2085***
	(0.0105)	(0.0167)	(0.0119)	(0.0333)
corporate	0.3093***	1.2208***	0.2897	-0.0157
	(0.0091)	(0.0144)	(0.0101)	(0.0319)
average_mun_age	0.0051***	-0.0010)	0.0053*	0.0202***
average_mun_area	-0.0003	0.0049***	-0.0013**	-0.0001
	(0.0004)	(0.0005)	(0.0004)	(0.0012)
	(0.0018)	(0.0025)	(0.0021)	(0.0061)
mun_pig/poultry_share	0.0104	-0.0373	0.0062	-0.1104
	(0.0354)	(0.0504)	(0.0407)	(0.1149)
mun_corporate_share	-0.0545	-0.4082***	-0.0202	-0.0499
	(0.0358)	(0.0497)	(0.0409)	(0.1140)
average_sar_age	0.0181**	0.0435***	0.026***	0.1549***
average_sar_area	0.0018***	0.0026**	0.0032***	-0.0022
	(0.0007)	(0.0010)	(0.0008)	(0.0024)
	(0.0071)	(0.0099)	(0.0081)	(0.0232)
sar_pig/poultry_share	0.1488**	1.0079***	0.1223	0.1305
	(0.0709)	(0.0993)	(0.0816)	(0.2306)
sar_corporate_share	0.1402**	0.4428***	0.1994**	0.0997
	(0.0677)	(0.0962)	(0.0770)	(0.2210)
unempl_rate	0.0104***	0.0311***	0.0164***	0.0797***
	(0.0033)	(0.0046)	(0.0038)	(0.0105)
Type shares		17.90%	54.20%	27.90%
Number of observations	317,177		317,177	
Correct predictions	92.73%		93.85%	
Log pseudo-likelihood	-76,323		-73,696	
AIC	152,684		147,470	
BIC	152,886%		147,886	
AIC3	152,703		147,509	

Note: *, ** and *** indicate significance at 10%, 5% and 1% levels, respectively; standard errors in parentheses.

Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

Table 3. Z-scores for testing the equality of the estimated farm types' coefficients from the mixture probit model for farm survival

Variable code	z-score (1-2)	p-value	z-score (1-3)	p-value	z-score (2-3)	p-value
intercept	-3.61	0.00	-44.24	0.00	-43.53	0.00
time trend	-1.22	0.22	12.49	0.00	13.50	0.00
age	7.18	0.00	98.55	0.00	97.84	0.00
age_square	-2.49	0.01	-107.10	0.00	-106.99	0.00
area	55.92	0.00	19.68	0.00	-9.86	0.00
area_square	-26.93	0.00	-1.59	0.11	11.02	0.00
agri_profit	67.28	0.00	22.27	0.00	10.43	0.00
agri_profit_×_retirement	11.67	0.00	-4.94	0.00	-7.58	0.00
pig/poultry	-8.75	0.00	0.03	0.98	5.11	0.00
corporate	-53.11	0.00	-35.38	0.00	-9.14	0.00
average_mun_age	1.90	0.06	3.21	0.00	2.31	0.02
average_mun_area	-8.95	0.00	-3.81	0.00	0.91	0.36
mun_pig/poultry_share	0.67	0.50	-0.58	0.56	-0.96	0.34
mun_corporate_share	6.03	0.00	2.88	0.00	-0.25	0.80
saa_age	-1.37	0.17	4.42	0.00	5.24	0.00
saa_area	0.52	0.60	-1.87	0.06	-2.19	0.03
saa_pig/poultry	-6.89	0.00	-3.49	0.00	0.03	0.98
saa_corporate	-1.98	0.05	-1.42	0.16	-0.43	0.67
unempl_rate	-2.45	0.01	4.25	0.00	5.68	0.00

Note: z-scores for testing the null hypothesis that coefficients of two different types are equal (Paternoster *et al.*, 1998).

Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

to survive is found for the majority of farms (about 54%) while the positive impact is observed only for about 18% of farms. Computed z-scores show that these opposite effects are significantly different at a 1% level (see Table 3). The different effects of neighbouring farms' size observed for the various groups may explain the insignificant impacts for the overall population, that is to say when such unobserved heterogeneity is not considered.

Referring to the discussion in Section 2, the two first types could mostly consist of business oriented farms where farm holders are mainly motivated by profit maximisation. The resulting negative impact of neighbouring farms' size on the probability to survive for type 2 may indicate that farms in this type are rather competitors for land, while the opposite effect for farms in type 1 may originate from positive spill overs of new technology adoption for these farms. Contrary to the two first farm types, the impact of neighbouring farms' size is highly non-significant for the third type which accounts for about 28% of the farm population in Brittany. Referring again to the discussion in Section 2, this initially unexpected type could comprise farms characterised by

Table 4. Main characteristics of each type of farms identified with the mixture probit model for farm survival

Variable code	Means			t-tests of equality of means		
	Type1	Type2	Type3	(1-2)	(1-3)	(2-3)
age	49.25	48.27	47.77	39.52	60.86	21.35
area	43.91	50.49	48.04	-63.66	-37.70	22.22
agri_profit	8.80	11.02	11.36	-77.33	-78.59	-9.54
pig/poultry	0.17	0.18	0.19	-17.96	-25.60	-9.20
corporate	0.43	0.48	0.45	-35.74	-11.45	21.80
average_mun_age	48.46	48.25	48.31	36.31	25.23	-9.63
average_mun_area	48.19	48.30	48.58	-3.16	-10.66	-7.85
mun_pig/poultry_share	0.18	0.18	0.19	-16.35	-21.09	-6.06
mun_corporate_share	0.46	0.46	0.47	8.35	-13.38	-21.35

Note: t-tests for the null hypothesis that the means of two different types are equal; the significance level are not reported since all t-statistics are significant at a level of 5%.

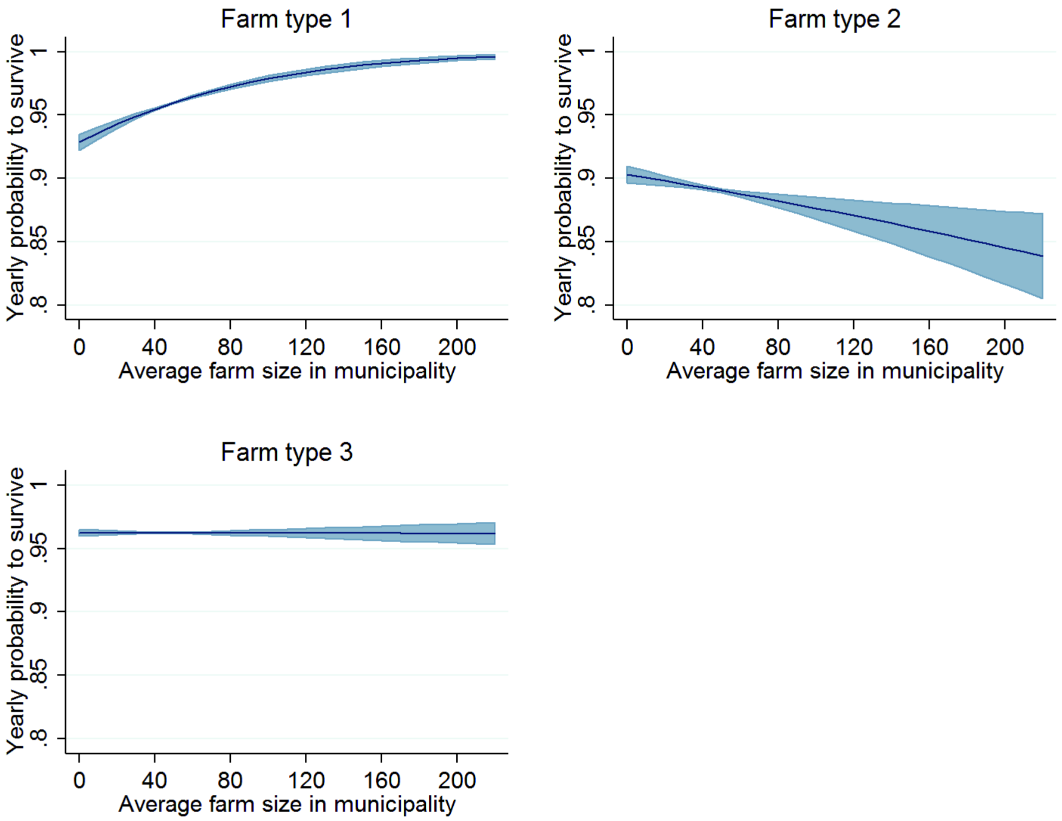
Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

prevailing non-pecuniary motives. It could be also the case of business oriented farms that have already reached their optimal economic size. The probability of survival for such farms may be therefore independent from the size of their neighbours. This result is in line with the impact of the average farm size at the small agricultural region, which has no significant effect on the probability to survive of this third farm type, contrary to the two first types. This suggests that the farming context has no specific influence on the persistence of such (third type) farms in the sector. This interpretation is confirmed by the positive impact of the time trend, meaning that the probability to survive increases for those (third type) farms over time, while the inverse trend is observed for farms that compete for land. This result is consistent with the evolution of farm size over the years: the larger the neighbours and the higher the competition for land, then the more difficult it becomes to innovate since new adoptions generally require more land.

The descriptive statistics for farm types are reported in Table 4. It should be noted that even if the differences between the means of some considered characteristics for two types are very small, the hypothesis of equality of means is rejected in all cases at the 0.1% level. This high level of significance could be due to the large size of the sample used for our estimation and should therefore be interpreted with due care.‡ Briefly, the results show that larger corporate farms are more likely to behave as competitors for land than farms in the other types. This result conforms with our discussion in Section 2 since such farms are more likely to be business oriented and thus mainly motivated by profit maximisation. Conversely, individual farmers with

‡The t-test statistics are computed as: $t = \frac{\bar{\mu}_1 - \bar{\mu}_2}{\sqrt{\left(s^2_{\mu_1}/n_{\mu_1} + s^2_{\mu_2}/n_{\mu_2}\right)}}$. As it can be seen from the formula, the larger the number of observations the larger the resulting t values.

Fig. 1. Probability to survive for varying municipality-level average farm sizes by unobserved farm types (predicted margins with 95% confidence intervals)



Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

smaller operated farm sizes are more likely to benefit from positive spill overs of new technology adoption.

Figure 1 reports the probability that an average farm remains active from year to year from 2003 to 2013 with respect to the average farm size calculated at the municipality level. Three panels are provided, one for each type of farms. The figure shows that, overall, the probability to survive is lower for competitors for land and this probability decreases with neighbouring farms' size (farm type 2). The opposite effect is observed for farms that benefit from positive spill overs of new technology adoption (farm type 1). Figure 1 also shows that the probability to survive is higher and does not vary with the neighbouring farms' size for farms having mainly non-pecuniary motives or already that reached their optimal size (farm type 3).

The impacts of own farm's and farmer's characteristics on the probability to survive also vary according to the specific type a farm belongs to. For example, the first type of farms is characterised by a negative impact of the age of farm holder and of the total agricultural profit, in contrast to the other types for which the impacts are positive. While the result regarding total agricultural profit is more difficult to interpret, the negative impact of age may be explained by the fact that young farmers may be more likely to adapt their production capacity using new technology. This result is consistent with the positive impact of the average age of farm holders at the municipality level on the probability for a farm to survive. The younger the neighbours, the more competitive they are because of a possible relative higher motivation or capacity to innovate by adopting a new technology.

In addition to the fact that the mixture probit model enables identifying specific impacts of neighbouring farm size, it presents some other advantages in comparison to the pooled estimation where unobserved heterogeneity is not considered. The results show that the finite mixture model performs better in terms of all criteria reported at the bottom of Table 2 (correct predictions, log-likelihood, and AIC, BIC, AIC3 information criteria). Furthermore, the finite mixture probit model is more accurate in predicting farm survival in Brittany. The superiority of the mixture model in particular comes from the specificity value. Indeed, the mixture model performs about 15% better in predicting farm exit in Brittany than the pooled estimation.

6.2. Farm growth model

Table 5 reports the estimated parameters for both the pooled estimation and the mixture of simple linear regressions for the farm growth process.

Here also, the results from the pooled estimation are conformed to our expectations. In particular, farm size and agricultural profit as well as corporate legal status have a positive impact on farm growth, while farmer's age has the opposite effect. Farms' characteristics at the small agricultural region level have the same impact as at the farm level, suggesting again spatial correlation of these characteristics, except for the legal status which negatively influences farm growth when the small agricultural region level is considered. This may indicate that spatial correlation exists at a higher level for the legal status of farms. Conversely, the results also show that neighbouring farms' characteristics at the municipality level have opposite effects on farm growth, except for neighbouring farms' size which positively affects farm growth. This suggests an indirect effect resulting from the negative impacts on farm survival, as shown in the previous section.

Turning to the mixture model, the farm population is divided into three types, characterised by specific impacts of the explanatory variables. However, in contrast to our expectation, the mixture linear regression model does not allow identifying opposite impacts of neighbouring farms' size on the own farm's operating size growth. Even though the results show that the studied farm population may be divided into more than one homogeneous type (according to the chosen information criteria), farm size growth is found to be positively related to neighbouring farms' size for all the endogenously determined homogeneous types. The positive impact of

Table 5. Estimated parameters for both the pooled and the mixture linear regression model for farm growth

Variable code	Pooled	Mixture		
		type 1	type 2	type 3
intercept	4.0651	-147.2279***	-59.9425***	-17.5979***
	(5.6072)	(1.0430)	(14.6292)	(1.0048)
time_trend	0.2524***	0.0197*	0.4053***	-0.0964***
	(0.0475)	(0.0087)	(0.1218)	(0.0087)
age	-0.711***	-0.0976***	0.0854	0.0025
	(0.0386)	(0.0110)	(0.1563)	(0.0066)
age_square	0.0003	0.0008***	-0.0093***	-0.0002***
	(0.0004)	(0.0001)	(0.0017)	(0.0001)
ln_area	5.6571***	52.6593***	4.8363***	9.4797***
	(0.0372)	(0.0431)	(0.0325)	(0.0176)
agri_profit	0.3538***	0.0040***	0.3598***	0.0100***
	(0.0100)	(0.0011)	(0.0151)	(0.0012)
agri_profit_×_retirement	0.1108***	-0.0138***	-0.1086***	0.0123***
	(0.0143)	(0.0017)	(0.0300)	(0.0030)
pig/poultry	-9.7902***	0.2651***	-18.2485***	0.1193***
	(0.1809)	(0.0317)	(0.4210)	(0.0355)
corporate	27.1756***	0.4472***	24.545***	0.9315***
	(0.1238)	(0.0234)	(0.3401)	(0.0366)
average_mun_age	0.7470***	0.0309***	1.3864***	0.0031
average_mun_area	0.8599***	0.0085***	1.2661***	0.0034**
	(0.0072)	(0.0011)	(0.0167)	(0.0012)
	(0.0306)	(0.0058)	(0.0816)	(0.0052)
mun_pig/poultry_share	12.267***	-0.5257***	21.3584***	0.9872***
	(0.6074)	(0.1031)	(1.5000)	(0.1132)
mun_corporate_share	-31.8797***	-0.3832***	-50.9605***	0.7515***
	(0.6083)	(0.1049)	(1.5183)	(0.1081)
average_saa_age	-0.4408***	-0.1048***	0.0917	0.1777***
average_saa_area	0.0585***	-0.0113***	0.0718*	0.0080***
	(0.0122)	(0.0023)	(0.0331)	(0.0021)
	(0.1172)	(0.0216)	(0.3054)	(0.0209)
saa_pig/poultry_share	-4.3395***	0.2592	-25.2777***	0.3612
	(1.2088)	(0.2054)	(2.9371)	(0.2179)
saa_corporate_share	-5.6855***	-0.8635***	8.6827**	0.2063
	(1.1420)	(0.2066)	(2.9053)	(0.2046)
unempl_rate	0.2425***	0.1176***	0.4492***	-0.0651***
	(0.0519)	(0.0095)	(0.1371)	(0.0104)
Lnsigma	3.4262***	1.3233***	3.7508***	1.1468***
	(0.0032)	(0.0038)	(0.0046)	(0.0039)
Type shares		43.72%	27.68%	28.60%
Number of observations	294,288		294,288	
Log pseudo-likelihood	-1,425,877		-1,036,676	
AIC	2,851,794		2,073,434	
BIC	2,852,005%		2,073,868	
AIC3	2,851,814		2,073,475	

Note: *, ** and *** indicate significance at 10%, 5% and 1% levels, respectively; standard errors in parentheses.

Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

Table 6. Z-scores for testing the equality of the estimated farm types' coefficients from the mixture simple linear regression for farm growth

Variable code	z-score (1-2)	p-value	z-score (1-3)	p-value	z-score (2-3)	p-value
intercept	5.95	0.00	89.51	0.00	2.89	0.00
time_trend	3.16	0.00	-9.39	0.00	-4.11	0.00
age	1.17	0.24	7.79	0.00	-0.53	0.60
age_square	-5.93	0.00	-8.20	0.00	5.31	0.00
ln_area	-886.48	0.00	-928.04	0.00	125.53	0.00
agri_profit	23.52	0.00	3.57	0.00	-23.11	0.00
agri_profit_×_retirement	-3.15	0.00	7.52	0.00	4.00	0.00
pig/poultry	-43.85	0.00	-3.06	0.00	43.47	0.00
corporate	70.69	0.00	11.14	0.00	-69.03	0.00
average_mun_age	16.57	0.00	-3.57	0.00	-16.92	0.00
average_mun_area	75.26	0.00	-3.13	0.00	-75.55	0.00
mun_pig/poultry_share	14.56	0.00	9.88	0.00	-13.54	0.00
mun_corporate_share	-33.23	0.00	7.54	0.00	33.97	0.00
saa_age	0.64	0.52	9.40	0.00	0.28	0.78
saa_area	2.51	0.01	6.32	0.00	-1.92	0.06
saa_pig/poultry	-8.67	0.00	0.34	0.73	8.71	0.00
saa_corporate	3.28	0.00	3.68	0.00	-2.91	0.00
unempl_rate	-2.41	0.02	-12.98	0.00	-3.74	0.00
Lnsigma	408.72	0.00	-32.40	0.00	-433.94	0.00

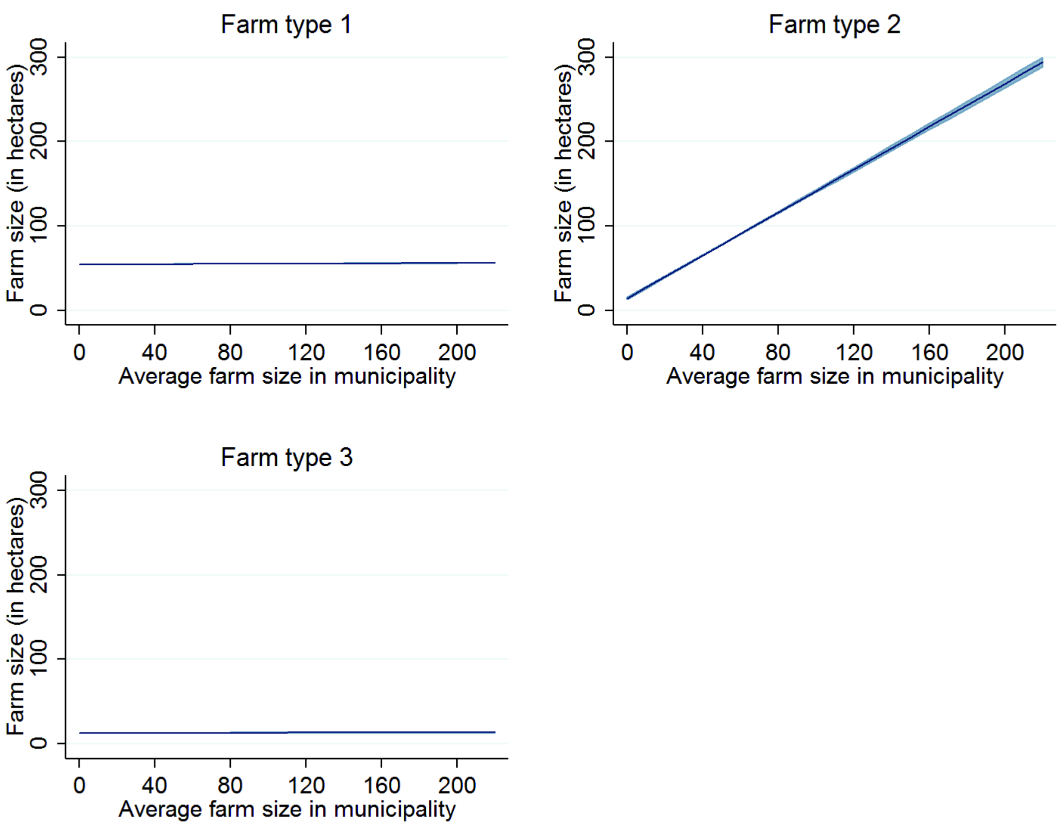
Note: z-scores for testing the null hypothesis that coefficients of two different types are equal (Paternoster *et al.*, 1998).

Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

neighbouring farms' size on own size growth whatever the farm type may indicate that farmers continually adjust their operated farm size during their lifespan in the sector. However, the magnitude of the impact of neighbouring farms' size differs from type to type, with the differences across types being significant at the level of 0.1% (see Table 6).

Figure 2 presents the evolution of farm size with respect to neighbouring farms' size for the three identified farm types. It shows that the impact of neighbouring farms' size is relatively low for farm types 1 and 3, while farm size varies almost proportionally to neighbouring farms' size in type 2. In this latter case, the impact of neighbouring farms' size is almost 150 times higher and more than 350 times higher than it is for type 1 and type 3, respectively. This may reveal that farms belonging to types 1 and 3 have lower pecuniary motives than those in type 2, or that they may have already reached their optimal size implying that they are less sensitive to the evolution of neighbouring farms' size.

Fig. 2. Total land used for varying average farm size in municipality by type of farms (predicted margins with 95% confidence intervals).



Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

Table 7. Main characteristics of types from the mixture of normal regression for farm growth

Variable code	Means			t-tests of equality of means		
	Type1	Type2	Type3	(1-2)	(1-3)	(2-3)
age	47.30	45.55	51.23	52.21	-97.30	-132.96
area	54.52	79.30	12.85	-130.84	749.65	-107.48
agri_profit	12.29	12.59	6.44	-6.11	109.40	100.63
pig/poultry	0.15	0.24	0.18	-56.042	-19.48	33.51
corporate	0.47	0.77	0.21	-154.65	134.21	287.64
average_mun_age	48.18	48.21	48.56	-2.49	-37.26	-32.49
average_mun_area	49.36	51.93	43.81	-46.60	98.91	-185.24
mun_pig/poultry_share	0.19	0.19	0.17	-14.62	36.04	46.55
mun_corporate_share	0.47	0.48	0.44	-23.20	45.84	62.99

Note: t-tests for the null hypothesis that the means of two different types are equal; the significance level are not reported since all t-statistics are significant at a level of 5%.

Source: MSA COTNS database, Bretagne 2004-2014 - authors' calculations

Descriptive statistics for each unobserved type are presented in Table 7. The figures show that farms in type 2 are on average much larger and more likely to be with a corporate legal form in comparison to types 1 and 3. Farms in type 2 are thus more likely to increase their operated size over time since they may have higher financial capacity. Conversely, farms belonging to type 3 are on average smaller and with older farm holders than those in the two other types, and are operated predominantly under individual legal status, with on average younger farmers as neighbours. This could explain why type 3 farms are more likely to decrease their operated size over time. As discussed in the previous section, older farmers may be less motivated to increase their farm size; in addition, they may face higher competition for land when surrounded by younger farmers. Furthermore, farms operated under individual legal status may face higher financial credit constraints.

7. Concluding remarks

The study conducted in this paper underlines the importance of accounting for unobserved farm heterogeneity in spatial interdependence between farms when analysing farm structural change. This was made possible by a modelling approach that enables endogenously grouping farms within specific homogeneous types. This approach allows identifying specific relationships between farms via the impact of neighbouring farms' size, measured at the municipality level, on farm survival and growth. The application to a panel of French farms located in Brittany shows that the relationship between farms in this region is rather in terms of competition for land than in terms of positive spill overs of new technology adoption. This results in a negative impact of neighbouring farms' size on the probability to survive for a majority of farms. However, for a about 18% of the farm population, the neighbouring farms' size has no significant impact on the probability to survive, suggesting the existence of potential non-pecuniary motives for these farms.

In contrast to the probability to survive, neighbouring farms' size is positively related to farm growth. Indeed, while three unobserved farm types are also endogenously identified in the case of farm growth, a positive effect of neighbouring farms' size is evidenced for all the three types of farms. This suggests that, even though neighbouring farms' size does not affect the probability of some farms to remain in business from year to year, farms in general tend to adjust their operated size over time. However, heterogeneity in the growth process is evidenced since the impact of neighbouring farms' size varies in magnitude according to farm types. However, descriptive statistics show that the resulting farm types from the growth model are more related to some observed characteristics than the farms types from the probit model. These results confirm that neighbouring farms' size may differently influence farm survival and growth, and suggest that farms should not be considered as isolated entities and that agricultural policies should take into account potential relationships between farms.

While this study clearly adds to the existing literature, the analysis could be improved in three different ways. Firstly, the impact of neighbouring farms' size is investigated here by using the average farm size at the municipality as a proxy. However, farms may compete for land in other municipalities in addition to their own municipality. Hence, investigating the impact of neighbouring farms' size using a spatial weighting matrix constructed at the municipality level or, if possible, at the farm level (using appropriate data sources that include the exact location of farms), could help estimate more efficiently the impact of neighbouring farms' characteristics. Secondly, the two models used for the present analysis (farm survival and farm growth) are estimated separately. However, it is clear that the survival and growth processes in farming are related to each other. Not accounting for this link may bias the estimation results. In this case, a sample selection model or a two-part model could be a way of improvement. Thirdly, some other factors, such as subsidies received by the farms and their neighbours, may have a significant impact on farm survival and growth as it has been shown by previous studies (see Storm *et al.* (2015) for a recent example). Including such variables in the analysis may thus improve the understanding of structural change in farming.

References

Allaire, G., Poméon, T., Maigné, E., Cahuzac, E., Simioni, M. and Desjeux, Y. (2015). Territorial analysis of the diffusion of organic farming in france: Between heterogeneity and spatial dependence. *Ecological Indicators* 59: 70 – 81.

Andrews, R. L. and Currim, I. S. (2003). Retention of latent segments in regression-based marketing models. *International Journal of Research in Marketing* 20: 315–321.

Aubert, M. and Perrier-Cornet, P. (2009). Is there a future for small farms in developed countries? evidence from the French case. *Agricultural Economics* 40: 797–806.

Bakucs, Z., Bojnec, S., Fertő, I. and Latruffe, L. (2013). Farm size and growth in field crop and dairy farms in France, Hungary and Slovenia. *Spanish Journal of Agricultural Research* 11: 869–881.

Bollman, R. D., Whitener, L. A. and Tung, F. L. (1995). Trends and patterns of agricultural structural change: A Canada–US comparison. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 43: 15–28.

Bowman, M. S. and Zilberman, D. (2013). Economic factors affecting diversified farming systems. *Ecology and Society* 18: 33.

Breustedt, G. and Glauben, T. (2007). Driving forces behind exiting from farming in Western Europe. *Journal of Agricultural Economics* 58: 115–127.

Case, A. (1992). Neighborhood influence and technological change. *Regional Science and Urban Economics* 22: 491–508.

Conradt, S., Bokusheva, R., Finger, R., Kussaiynov, T. *et al.* (2014). Yield trend estimation in the presence of farm heterogeneity and non-linear technological change. *Quarterly Journal of International Agriculture* 53: 121–140.

Daniels, T. L. (1986). Hobby farming in America: Rural development or threat to commercial agriculture? *Journal of Rural Studies* 2: 31–40.

Dong, F., Hennessy, D. A. and Jensen, H. H. (2010). Contract and exit decisions in finisher hog production. *American Journal of Agricultural Economics* 92: 667–684.

Eastwood, R., Lipton, M. and Newell, A. (2010). Farm size. *Handbook of agricultural economics* 4: 3323–3397.

Gaté, R. and Latruffe, L. (2016). Difficultés rencontrées lors de la transmission d'exploitations agricoles. le cas de la bretagne. *Économie rurale* 1: 5–24.

Goddard, E., Weersink, A., Chen, K. and Turvey, C. G. (1993). Economics of Structural Change in Agriculture. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 41: 475–489.

Greene, W. (2004). Convenient estimators for the panel probit model: Further results. *Empirical Economics* 29: 21–47.

Greene, W. H. (2012). *Econometric analysis (International edition)*. Pearson, 7th ed.

Hallam, A. (1991). Economies of size and scale in agriculture: An interpretive review of empirical measurement. *Review of Agricultural Economics* 13: 155–172.

Harrington, D. H. and Reinsel, R. D. (1995). A synthesis of forces driving structural change. *Canadian Journal of Agricultural Economics* 43: 3–14.

Hess, S., Bierlaire, M. and Polak, J. (2006). A systematic comparison of continuous and discrete mixture models. Tech. rep.

Holloway, G., Lacombe, D. and LeSage, J. P. (2007). Spatial econometric issues for bio-economic and land-use modelling. *Journal of Agricultural Economics* 58: 549–588.

Holloway, G., Shankar, B. and Rahmanb, S. (2002). Bayesian spatial probit estimation: a primer and an application to hyv rice adoption. *Agricultural Economics* 27: 383–402.

Holloway, L. (2002). Smallholding, hobby-farming, and commercial farming: Ethical identities and the production of farming spaces. *Environment and Planning A* 34: 2055–2070.

Howley, P. (2015). The Happy Farmer: The Effect of Nonpecuniary Benefits on Behavior. *American Journal of Agricultural Economics* 97: 1072 –1086.

Howley, P., Dillon, E. and Hennessy, T. (2014). It's not all about the money: understanding farmers' labor allocation choices. *Agriculture and Human Values* 31: 261–271.

Jackson-Smith, D. B. (1999). Understanding the Microdynamics of Farm Structural Change: Entry, Exit, and Restructuring among Wisconsin Family Farmers in the 1980s. *Rural Sociology* 64: 66–91.

Kyriazidou, E. (1997). Estimation of a panel data sample selection model. *Econometrica* 65: 1335–1364.

Latruffe, L. and Piet, L. (2014). Does land fragmentation affect farm performance? a case study from brittany, france. *Agricultural Systems* 129: 68–80.

Maybery, D., Crase, L. and Gullifer, C. (2005). Categorising farming values as economic, conservation and lifestyle. *Journal of Economic Psychology* 26: 59 – 72.

McLachlan, G. and Krishnan, T. (2007). *The EM algorithm and extensions*, *382*. John Wiley & Sons.

McLachlan, G. and Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.

Mzoughi, N. (2011). Farmers adoption of integrated crop protection and organic farming: Do moral and social concerns matter? *Ecological Economics* 70: 1536 – 1545.

Paternoster, R., Brame, R., Mazerolle, P. and Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology* 36: 859–866.

Pennings, J. M. and Garcia, P. (2004). Hedging behavior in small and medium-sized enterprises: The role of unobserved heterogeneity. *Journal of Banking & Finance* 28: 951–978.

Piet, L. and Cariou, S. (2014). Le morcellement des exploitations agricoles françaises. *Économie rurale* : 107–120.

Shapiro, D., Bollman, R. D. and Ehrensaft, P. (1987). Farm size and growth in canada. *American Journal of Agricultural Economics* 69: 477–483.

Storm, H. and Heckelei, T. (2016). Direct and indirect spatial interaction of Norwegian farm growth. Selected paper presented at the AAEA annual meeting. July 31 - August 2, Boston (USA).

Storm, H., Mittenzwei, K. and Heckelei, T. (2015). Direct Payments, Spatial Competition, and Farm Survival in Norway. *American Journal of Agricultural Economics* 97: 1192–1205.

Sumner, D. A. and Leiby, J. D. (1987). An econometric analysis of the effects of human capital on size and growth among dairy farms. *American Journal of Agricultural Economics* 69: 465–470.

Teillard, F., Allaire, G., Cahuzac, E., Léger, F., Maigné, E. and Tichit, M. (2012). A novel method for mapping agricultural intensity reveals its spatial aggregation: Implications for conservation policies. *Agriculture, Ecosystems & Environment* 149: 135 – 143.

Train, K. E. (2008). Em algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling* 1: 40–69.

Trujillo-Barrera, A., Pennings, J. M. E. and Hofenk, D. (2016). Understanding producers' motives for adopting sustainable practices: the role of expected rewards, risk perception and risk tolerance. *European Review of Agricultural Economics* 43: 359 –382.

Weiss, C. R. (1999). Farm Growth and Survival: Econometric Evidence for Individual Farms in Upper Austria. *American Journal of Agricultural Economics* 81: 103–116.

Willock, J., Deary, I. J., Edwards-Jones, G., Gibson, G. J., McGregor, M. J., Sutherland, A., Dent, J. B., Morgan, O. and Grieve, R. (1999). The Role of Attitudes and Objectives in Farmer Decision Making: Business and Environmentally-Oriented Behaviour in Scotland. *Journal of Agricultural Economics* 50: 286–303.