Farm segmentation and agricultural policy impacts on structural change: evidence from France

Legrand D. F. SAINT-CYR
AGROCAMPUS OUEST, UMR1302 SMART, F-35000 Rennes, France
legrand.saintcyr@agrocampus-ouest.fr

Paper prepared for presentation at the 149th EAAE Seminar 'Structural change in agri-food chains: new relations between farm sector, food industry and retail sector'
Rennes, France, October 27-28, 2016

Abstract

This study aims at investigating the impact of agricultural policies on structural change in framing. As not all farms may behave alike, a mixed Markov chain modelling approach is applied to capture for unobserved heterogeneity transition process of farms. A multinomial logit specification is used for transition probabilities and the parameters are estimated by maximum likelihood method and the expectation-maximization (EM) algorithm. An empirical application to an unbalanced panel from 2000 to 2013 shows that the French farms mainly consists of a mixture two farm types characterized by specific transition processes. The main result of this paper is that the impact of farm subsidies from both the two pillars of the Common Agricultural policy (CAP) highly depends on the type membership of farms. From this result it is argued that more attention should be paid to both observed and unobserved farm heterogeneity in assessing agricultural policy impact on structural change in farming.

1. Introduction

The farming sector has faced important structural change over recent decades. In particular, the number of farms has decreased sharply and their average size increasing continually in most developed countries, implying some changes in farm size distribution. These changes may have important consequences for equity among farmers, productivity and efficiency of farming (Weiss, 1999). Structural change has been therefore the subject of considerable interest to agricultural economists and policy makers. Many theoretical studies pointed out potential impacts of agricultural policy on farm size changes but, these impacts remain ambiguous (see for example Goddard *et al.* (1993); Harrington and Reinsel (1995)). Since then, several studies empirically investigated the impact of some recent policy programmes on structural change in various farming contexts. This study aims at investigating the impacts of farm subsidies from the Common agricultural Policy (CAP) on farm structural change especially in France.

It has become quite common in agricultural economics to study structural change in farming and the impact of time-varying variables including agricultural policy by applying the so-called Markov chain model (MCM). Focusing on farm number and size evolution, Markov chain has been proved to be a convenient modelling approach to represent transition process of farms across category of sizes (Bostwick, 1962; Padberg, 1962; Krenz, 1964). Basically, this model states that farm size at a given time period is the result of a probabilistic process where future farm size only depends on size in the immediately previous period. In general, a first order process is assumed. Recently, Stokes (2006) showed that a Markovian transition process may derived from a structural model of inter-temporal profit maximization, given theoretical grounds to using the MCM. This paper add to the existing literature mainly in two ways. Firstly, a mixture modelling framework is applied to take into account potential unobserved farm heterogeneity in the analysis of structural change in farming. Secondly, transition probabilities are specified at individual farm level within a discrete choice modelling approach to analyse transition process at farm level.

In agricultural economics, heterogeneity issue is mostly left in the background in the analysis of farm structural change may be because previous studies generally focus on specific farm types. However, heterogeneity may be crucial to understand structural change process in farming because it results from individual farmers' decisions (Freshwater and Reimer, 1995). Farm heterogeneity may originate from several sources. One of the most important sources of farm heterogeneity is farmers' motivation. While farmers are normally supposed to maximise their total profit from farming activities, it has been shown that not all farmers give a priority to profit maximisation (Maybery *et al.*, 2005; Mzoughi, 2011; Howley *et al.*, 2014). It is the case of environmentally oriented farms (Willock *et al.*, 1999) or some hobby farms (Daniels, 1986; Holloway, 2002). The existence of non-financial/pecuniary motives or potential farming lifestyle values may shape farmers' behaviours (Hallam, 1991; Harrington and Reinsel, 1995; Howley, 2015). The ability to change operated farm size may also depend on some other factors such as accessibility to inputs (land, new technology), managerial capacity, risk perception, risk tolerance of farmers, etc. (Bowman and Zilberman, 2013; Conradt *et al.*, 2014; Trujillo-Barrera *et al.*, 2016). Therefore, the implicit homogeneity assumption of the usual MCM, i.e all farms have the same probability to change category of sizes, does not hold.

Previous studies tried to control for farm heterogeneity in modelling structural change in non-stationary MCM. However, only observed sources of heterogeneity has been considered so far (see Zimmermann and Heckelei (2012) for a recent example). These approaches seem to be restricted since not all sources of farm heterogeneity are observable or can be linked to observed farm and farmer characteristics. The mixed Markov chain model (M-MCM) applied here captures unobserved heterogeneity in the transition process of farms. The existence of different transition processes in farming may reflect heterogeneity in farmers' behaviour, which may related both to observed and unobserved farm and farmer characteristics.

To the best of our knowledge, Saint-Cyr and Piet (2014) are the first to evidence heterogeneity in farm transition process. As in other strands of economics, the authors showed that even a restricted mover-stayer model represents more efficiently farm size dynamics than the usual MCM. This paper extends Saint-Cyr and Piet (2014)' approach by: firstly, allowing for more than two types of farms and also relaxing the 'pure stayer' assumption; secondly, by developing a non-stationary approach to study the impact of agricultural policy on farm size dynamics. This approach lead to separate the farm population into meaningful clusters with different transition patterns (Vermunt, 2010). As agricultural policy impacts may depend on some observed and unobserved farm and farmer characteristics, a mixture approach may therefore provide more informative results in the analysis of structural change.

In general, structural change is investigated using aggregate data, i.e cross-sectional observations of farm distribution into a finite number of size categories, because such data are most often easier to obtain than individual-level data. Structural change determinants are thus usually investigated at macro level. Focusing on agricultural policies, impacts of public supports on transition probabilities of farms across category of sizes are generally investigated using transition probability matrices estimated for the overall population of farms (see Huettel and Jongeneel (2011); Zimmermann and Heckelei (2012); Ben Arfa *et al.* (2015) for recent examples). Even using individual level data, yearly transition probability matrices are computed for the overall population of farms first; then, effects of exogenous variables are estimated on transition probabilities (see for example Rahelizatovo and Gillespie (1999)). In doing so, it's become difficult to take into account individual effects of farms. The discrete choice approach adopted here to model farm transitions enables to more easily incorporate individual farm effects on structural change.

This paper is structured as follows. Section 2 presents the proposed model, the methods of specification of transition probabilities and the estimation procedure. In section 3, methods for assessing the model are presented. Section 4 reports the application to a panel of French farms, starting with a description of the data used and explanatory variables investigated following by a presentation of the main results. Finally, concluding remarks are provided with some considerations on possible improvements of this study for further research.

2. The mixed Markov chain model (M-MCM)

Let by N the total number of farms in the population and K the total number of farm size categories (choice alternatives). As farm category of sizes are observed at discrete times, generally 1-year interval, a discrete-time process is assumed. Denote by y_{it} the category of sizes of a specific farm i ($i \in N$) at time t ($1 \leq t \leq T$). The indicator $y_{i1} = j$ ($\forall j = 0, 1, 2, \cdots, K$) if farm i is in category j at time $t = 1$. The category $j = 0$ may indicate entry in or exit from farming. Since farms may enter and leave the farming sector at different time points, the length of the vector $\mathbf{y}_i$ may vary across farms (*i.e*, $T_i \leq T$). Over time period T_i, the size evolution of a specific farm i can be represented by the vector $\mathbf{y}_i = (y_{i0}, y_{i1}, \cdots, y_{iT_i})$, where each element of $\mathbf{y}_i$ indicates farm category of sizes at each time point t ($t \in T_i$). As structural change in farming results from individual movements of farms across category of sizes, this process can thus be described using individual farms' movements.

In agricultural economics, transition process is generally supposed to follow a first order Markov process specially in the context of farm size changes over time (Zimmermann *et al.*, 2009). It is thus assumed that farm category of sizes at any time t (y_{it}) only depends on its immediately previous location, i.e its category of sizes at time $t-1$ (y_{it-1}). The Markov assumption implies that the observed random variables ($y_{i1}, y_{i2}, \cdots, y_{iT_i}$) are not independent from each other and $\mathbf{y}_i$ can be described by the probability function (Dias and Willekens, 2005):

$$f(\mathbf{y}_i) = \prod_{t=1}^{T_i} P(y_{it}|y_{it-1}) \tag{1}$$

where $P(y_{it}|y_{it-1})$ is the probability that farm i chooses a specific category of sizes at time t given its location at time $t-1$, so-called transition probability.

Suppose now that the observed random sample of farms is divided into G homogeneous types instead of just one, each type gathers farms with similar transition process. The density function of $\mathbf{y}_i$ as a discrete mixing distribution with G support points can thus be rewritten (McLachlan and Peel, 2004):

$$f(\mathbf{y}_i) = \sum_{g=1}^{G} \pi_g f_g(\mathbf{y}_i) \tag{2}$$

where $f_g(\mathbf{y}_i)$ is the probability function describing farm size dynamics in type g as specified in equation (1); and π_g, the mixing proportions, are non-negative and sum up to one. In statistics, π_g is called the mixing distribution and $f_g(\mathbf{y}_i)$ is called the mixed function (Train, 2009). As we defined a finite number of farm types, the mixed model can be also called a 'latent class model' with G latent transition processes. The density function of $\mathbf{y}_i$ is thus conditional on the mixing distribution and we can represent farm size dynamics as (Vermunt, 2010):

$$f(\mathbf{y}_i) = \sum_{g=1}^{G} P(g_i = g) \left[\prod_{t=1}^{T_i} P(y_{it} = k | y_{it-1} = j, g_i = g) \right] \tag{3}$$

From the above equation, it can be seen that under M-MCM farm size evolution has thus two set of probabilities. The first term are probabilities that farm i belongs to a specific farm type g while the second term are probabilities of making transitions across category of sizes given farm i belongs to type g. Both part of equation (3) can be specified as a function of exogenous variables. A semi parametric approach is applied here. Only transition probabilities are specified to study agricultural policy impacts on structural change.

2.1. *Specifying transition probabilities*

As farm category of sizes are mutually exclusive, finite and exhaustive, a discrete-choice approach is used to specify transition probabilities. The discrete choice approach assumes that farmers' choice of initial category of sizes as well as to make some transitions across categories can be represented by a random utility model (Train, 2009). Farmer's utility may represent in our case the net benefit that arise from choosing (or moving to) a specific category of sizes given its preceding location.

Denote by U_{ijkt} the utility of farm i arising upon moving from category of sizes j to another one k at time t. Under the basic behavioural assumption, it is supposed that farmers choose the category which maximized their utility. Therefore, a move from j to k (*i.e*, $y_{it} = k|y_{it-1} = j$) will be observed if and only if $U_{ijkt} \geq U_{ijlt}$ $(\forall j, k, l \in K)$. Farms staying in the same category two consecutive times is considered to make a transition from j to j. Under a mixture assumption, farm utility level is conditional on its type specific g. A farm belonging to a specific type g will thus make a transition from a specific category of sizes j to another one k if and only if $U_{ijkt|g} \geq U_{ijlt|g}$ $(\forall g \in G)$, where $U_{ijkt|g}$ is the utility of farms given belonging to the specific type g.

As this study is interested in determining the impacts of agricultural policy on farm size change, transition probabilities (*i.e*, the utility arising from moving across category of sizes) are specified as a function of some public supports and other causative factors. Under mixture assumption, the utility that would accrue to farm i upon moving from category j to another one k at time t given belonging to type g can be expressed as:

$$\begin{aligned} P(y_{it} = k|y_{it-1} = j, g_i = g, \mathbf{x}_{it-1}) =& P(U_{ijkt|g} \geq U_{ijlt|g}) \\ p_{ijkt|g} =& f(\mathbf{x}_{it-1}, \boldsymbol{\beta}_g, \epsilon_{ijkt|g}), \quad \forall t \in T_i \quad j, k, l \in K \quad g \in G \end{aligned} \tag{4}$$

where $\mathbf{x}_{it-1}$ are explanatory variables; $\boldsymbol{\beta}_g$ and $\epsilon_{ijkt|g}$ are respectively parameters to estimate and an *iid* random error terms specific to farm type-g. Explanatory variables are lagged 1-year since farmers' decisions for entering or the farming sector as well as for expansion, or contraction are likely dependent upon information available during the previous period.

Because farmers may face multiple choices at each occasion, it is econometrically convenient to used a multinomial specification (Greene, 2006). Assuming that the error terms $\epsilon_{ijkt|g}$ are randomly drawn from a Gumbel distribution (type I extreme value), the conditional probability of making a transition from a specific category of sizes j to the category k at time t is given by:

$$P(y_{it} = k|y_{it-1} = j, g_i = g, \mathbf{x}_{it-1}) = \frac{\exp(\boldsymbol{\beta}'_{jk|g}\mathbf{x}_{it-1})}{\sum_{l=1}^{K} \exp(\boldsymbol{\beta}'_{jl|g}\mathbf{x}_{it-1})} \quad (\forall j, k = 1, 2, \cdots, K) \tag{5}$$

where $\boldsymbol{\beta}_{jk|g}$ is a vector of parameters specific to each type of farm g and each transition from specific category j to another one k. Assuming permanence in the same category of sizes two consecutive years as the reference leads to state $\boldsymbol{\beta}_{jj|g} = \boldsymbol{0}$ $\forall g = 1, 2, \cdots, G$ and $\forall j = 1, 2, \cdots, K$ for identification.

2.2. *Estimation procedure*

The parameters of the model are estimated using the maximum likelihood estimation method. Let $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n)$ the observed random sample of farms obtained from the mixture density, where the vector $\mathbf{y}_i = (y_{i0}, y_{i1}, \cdots, y_{iT_i})$ gathering farm i locations over time period $T \geq T_i$. According to equation (3) and the model specification, state that:

$$P(y_{it} = k | y_{it-1} = j, g_i = g, \mathbf{x}_{it-1}) = p_{ijkt|g} = P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g})$$

Under a mixture assumption, the log-likelihood (LL) function for the parameters ($\boldsymbol{\beta}$) of the model, conditional on observing $\mathbf{y}$, writes:

$$LL(\boldsymbol{\beta}) = \sum_{i=1}^{N} \ln \left\{ \sum_{g=1}^{G} \pi_g \prod_{t=1}^{T_i} \prod_{j,k}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right]^{d_{ijkt}} \right\} \tag{6}$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_G)$ is a matrix of parameters with $\boldsymbol{\beta}_g = \left\{ \boldsymbol{\beta}_{jk|g} \right\} \forall g \in G$ and $j, k = 1, 2, \cdots, K$; the indicators $d_{ijkt} = 1$ if farm i moves from category of sizes j to category k at time t (*i.e*, $y_{it} = k | y_{it-1} = j$) and zero otherwise. Since farm type is unknown beforehand and given some numerical difficulties associated to the maximization of the above expression, the expectation-maximisation (EM) algorithm is generally used to estimate the parameters of such a model (McLachlan and Krishnan, 2007). The EM algorithm developed by Dempster *et al.* (1977) simplifies the complex log-likelihood in equation (6) in a set easily solvable log-likelihood functions by introducing a so-called 'missing variable'.

Let v_{ig} be a discrete unobserved variable indicating the type membership of each farm. The random vector $\mathbf{v}_i = (v_{i1}, v_{i2}, \cdots, v_{iG})$ is thus g-dimensional with $v_{ig} = 1$ if farm i belongs to type g and zero otherwise. Assuming that v_{ig} is unconditionally multinomial distributed with probability π_g, the complete likelihood for $(\boldsymbol{\beta}, \boldsymbol{\pi})$, conditional on observing $\mathbf{y}_c = (\mathbf{y}, \mathbf{v})$, therefore writes:

$$L_c(\boldsymbol{\beta}, \boldsymbol{\pi}) = \prod_{i=1}^{N} \prod_{g=1}^{G} \left\{ \pi_g \prod_{t=1}^{T_i} \prod_{j,k}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right]^{d_{ijkt}} \right\}^{v_{ig}} \tag{7}$$

where $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_G)$ vector gathering shares of farm type to also estimate. The complete log-likelihood is thus obtained as:

$$LL_c(\boldsymbol{\beta}, \boldsymbol{\pi}) = \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig} \ln \left\{ \pi_g \prod_{t=1}^{T_i} \prod_{j,k}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right]^{d_{ijkt}} \right\} \tag{8}$$

In this case, v_{ig} is called the 'posterior' probability that farm i belongs to the g-th type with $\mathbf{y}_i$ have been observed, that is $P(v_{ig} = 1 | \mathbf{y}_i)$, while π_g is a 'prior' probability of the mixture McLachlan and Peel (2004). This log-likelihood can be then divided into two components:

$$\begin{aligned} LL_1 &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig} \ln \pi_g \\ LL_2 &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \ln \left\{ P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right\} \end{aligned} \tag{9}$$

As the farm type is not observed, the posterior probability that farm i belongs to type g has to be estimated from the observations. The EM algorithm therefore consists in the four following steps:

(i) Initialization: Arbitrarily choose initial values $\mathbf{\Phi}^0 = (\boldsymbol{\phi}_1^0, \boldsymbol{\phi}_2^0, \cdots, \boldsymbol{\phi}_G^0)$ where $\boldsymbol{\phi}_g^0 = (\pi_g^0, \boldsymbol{\beta}_{jk|g}^0)$ $\forall j, k = 1, 2, \cdots, K$ and $\forall g = 1, 2, \cdots, G$ for the parameters of the model, with some parameters set to zero for identification as previously mentioned in section 2.1.

(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 parameters $\mathbf{\Phi}^p$. This conditional expectation probability, that is, the posterior probability $v_{ig}^{(p+1)} = v_{ig}(\mathbf{y}_i; \mathbf{\Phi}^p)$, can be obtained according to the Bayes' law:

$$v_{ig}^{(p+1)} = \frac{\pi_g^p \prod_{t=1}^{T_i} \prod_{j,k}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}^p) \right]^{d_{ijkt}}}{\sum_{h=1}^{G} \pi_g^p \prod_{t=1}^{T_i} \prod_{j,k}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|h}^p) \right]^{d_{ijkt}}} \tag{10}$$

Replacing v_{gi} by its expected value in equation (8) leads to the conditional expectation of the complete data log-likelihood.

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

$$\boldsymbol{\beta}^{(p+1)} = argmax_{\boldsymbol{\beta}} \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \ln \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right] \tag{11}$$

The maximization process of the above equation is straightforward. The transition probability parameters ($\hat{\boldsymbol{\beta}}^p$) are updated considering $v_{gi}(\mathbf{y}_i; \mathbf{\Phi}^p)$ as a weighted factor for each observation (Pacifico and Yoo, 2012). Then, the posterior probabilities of belonging to a specific type are updated as follows :

$$\pi_g^{(p+1)} = \frac{\sum_{i=1}^{N} v_{ig}^{(p+1)}}{\sum_{i=1}^{N} \sum_{h=1}^{G} v_{ih}^{(p+1)}}, \qquad \forall g \in G \tag{12}$$

(iv) Iteration: Return to expectation step (ii) using $\boldsymbol{\pi}^{(p+1)}$ and $\boldsymbol{\beta}^{(p+1)}$ and iterate until convergence of the observed log-likelihood given by equation (6). At convergence, the resulting parameters are considered as the optimal values ($\hat{\mathbf{\Phi}}$).

A problem which often occurs in a mixture analysis with several components is that some solutions may be suboptimal. Indeed, the non-concavity of the log-likelihood function in equation (6) does not allow the identification of a global maximum in the mixture model, even for discrete mixtures of multinomial logit (Hess *et al.*, 2006). Given the potential presence of a high number of local maxima, the EM solutions may be highly dependent on the initial values of $\mathbf{\Phi}^0$. Various techniques are used in the literature to avoid suboptimal solutions (see Baudry and Celeux (2015) for a short review). In this study, the EM algorithm are run with various initial values of parameters (randomly chosen) and the starting values providing the largest likelihood at convergence are chosen as the best ones.

3. Model assessment and elasticities

3.1. Choosing optimal number of farm types

The total number of components for mixture model can be chosen either by a priori assumptions or via information criteria. In the latter case, selection criteria are generally based on the value of $-2LL_G(\mathbf{y};\hat{\mathbf{\Phi}})$ of the model, where matrix $\hat{\mathbf{\Phi}}$ represents the maximum likelihood estimates adjusted for the number of free parameters in the model with 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 chosen (Andrews and Currim, 2003). The selection criteria are derived from the following formula:

$$C_G = -2\left\{LL_G(\mathbf{y};\hat{\mathbf{\Phi}})\right\} + \kappa N_G \tag{13}$$

where $LL_G(\mathbf{y};\hat{\mathbf{\Phi}})$ 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 a penalty constant. Different values of κ lead to the two well known information criteria: Akaike Information Criterion (AIC) with $\kappa = 2$ and the Bayesian Information Criterion (BIC) using $\kappa = \log N$ with N the total number of observations. Other information criteria can be also derived such as the Consistent Akaike Information Criterion (CAIC) stating $\kappa = \log N + 1$ and the modified AIC (AIC3) which uses $\kappa = 3$ as penalizing factor (Andrews and Currim, 2003; Dias and Willekens, 2005). For these heuristic criteria, smaller values mean more parsimonious models.

There is no general consensus in the literature for using a specific type of criteria to choose an optimal number of components for a mixture model. However, some studies suggest that the CAIC and AIC3 may be more useful in the context of mixture models since these criteria more severely penalize the addition of parameters (Andrews and Currim, 2003; Dias and Willekens, 2005).

3.2. *Probability elasticities*

The model tests whether the investigated exogenous variables have significant impacts on farm transition probabilities. As the estimated coefficients indicate marginal effects on the log-odds ratios of transition probabilities, the impacts of the explanatory variables are difficult to interpret directly (Greene, 2006). In this case, the impacts of explanatory variables are usually evaluate in terms of elasticities. The probability elasticities measure the effect of a 1% change in the ith explanatory variable (Zepeda, 1995). In the mixture model these probability elasticities may depend on farm type. Yearly transition probability elasticities for farms belonging to a specific type g are obtained as:

$$\boldsymbol{\delta}_{jkt|g} = \frac{\partial p_{jkt|g}}{\partial \mathbf{x}_{t-1}} \times \frac{\mathbf{x}_{t-1}}{p_{jkt|g}}, \qquad \forall j,k \in K \quad \forall g \in G \tag{14}$$

where $\boldsymbol{\delta}_{jkt|g}$ is a vector gathering elasticities at the means of the explanatory variables in vector $\mathbf{x}_{t-1}$; and $p_{jkt|g}$ is the probability to move from category j to category k at time period t given belonging to type g. The first term of equation (14) thus represents the marginal effects of explanatory variables and is given by (Greene, 2006):

$$\frac{\partial p_{jkt|g}}{\partial \mathbf{x}_{t-1}} = p_{jkt|g}\left(\boldsymbol{\beta}_{jk|g} - \sum_{l=1}^{K} \boldsymbol{\beta}_{jl|g} p_{jlt|g}\right) \tag{15}$$

where $\boldsymbol{\beta}_{jk|g}$ is the vector of estimated parameters.

By replacing the marginal effects in equation (14) leads to express the transition probability elasticities as:

$$\boldsymbol{\delta}_{jkt|g} = \left(\boldsymbol{\beta}_{jk|g} - \sum_{l=1}^{K} \boldsymbol{\beta}_{jl|g} p_{jlt|g}\right)\mathbf{x}_{t-1} \tag{16}$$

Given the constraint $\boldsymbol{\beta}_{jj|g} = 0$ for identification, the probability elasticities for the reference pair of transitions jj is thus obtained as:

$$\boldsymbol{\delta}_{jjt|g} = \left(- \sum_{l=1}^{K} \boldsymbol{\beta}_{jl|g} p_{jlt|g} \right) \mathbf{x}_{t-1} \tag{17}$$

3.3. Farm structure elasticities

Yearly structure elasticities are also derived to measure agricultural policy impacts on the distribution of farms across category of sizes. Farm structure elasticities measure the percentage change in the total number of farms in a specific category j at time t for a 1% change in the investigated explanatory variable (Zepeda, 1995).

Under the mixture modelling framework, the total number of farms in a specific category k at time t can be obtained as:

$$n_{kt} = \sum_{j=1}^{G} \pi_g \sum_{j}^{K} n_{jt-1} p_{jkt|g}, \qquad \forall k \in K, \quad \forall t \in T \tag{18}$$

where π_g is the probability of belonging to type g; n_{jt-1} is the total number of farms located in category of sizes j at time $t-1$; and $p_{jkt|g}$ the probability for farm i to make a transition from category of sizes j to category k at time t. Farm structure elasticities are then given by:

$$\boldsymbol{\eta}_{kt} = \frac{\partial n_{kt}}{\partial \mathbf{x}_{t-1}} \times \frac{\mathbf{x}_{t-1}}{n_{kt}} \tag{19}$$

Only transition probabilities in equation (18) depend on exogenous variables ($\mathbf{x}_{t-1}$). Farm structure elasticities can be therefore obtained using the corresponding probability marginal effects in equation (15). At any specific time t, farm structure elasticities are then derived as:

$$\boldsymbol{\eta}_{kt} = \left(\sum_{g=1}^{G} \pi_g \sum_{j=1}^{K} n_{jt-1} \frac{\partial p_{jkt|g}}{\partial \mathbf{x}_{t-1}} \right) \frac{\mathbf{x}_{t-1}}{n_{kt}} \tag{20}$$

where the marginal effects at the means of the corresponding explanatory variable are replaced by their values.

4. Empirical application

4.1. Data

For the empirical application, an unbalanced panel from the "Réseau d'Information Comptable Agricole" (RICA) database is used. RICA is the French implementation of the Farm Accountancy Data Network (FADN) and data are available from 2000 to 2013. FADN is an annual survey which is defined at the European Union (EU) level and is carried out in each member state. The information collected at the individual level relates to both the physical and structural characteristics of farms and their economic and financial characteristics. It is the only database providing information about the total subsidies received by farms from the Common Agricultural Policy (CAP) (see `http://ec.europa.eu/agriculture/rica/index.cfm` to learn more about FADN). In France, RICA is produced and disseminated by the statistical and foresight office of the French ministry for agriculture. It focuses on 'medium and large' farms and constitutes a stratified and rotating panel of approximately 7,000 farms surveyed each year. Some 10% of the sample is renewed every year so that, on average, farms are observed during 5 consecutive years. However, some farms may be observed only once, and others several, yet not consecutive, times. Some farms remained in the database over the whole of the studied period, i.e. fourteen consecutive years. Each farm in the dataset is assigned a weighting factor which reflects its stratified sampling probability, allowing for extrapolation at the population level (see `http://www.agreste.agriculture.gouv.fr/` to learn more about RICA France).

The study concentrates on farm size as defined in economic terms in order to consider all farms in the sample whatever their type of production. In accordance with the EU regulation (CE) N$^{\underline{o}}$1242/2008, European farms are classified into fourteen economic size (ES) categories, evaluated in terms of total standard output (SO) expressed in Euros. As mentioned before, in France, RICA focuses on 'medium and large' farms, those whose SO is greater than or equal to 25,000 Euros; this corresponds to ES category 6 and above. According to the EU regulation (CE) N$^{\underline{o}}$1242/2008, the nine ES categories available in RICA are aggregated into three categories: strictly less than 100,000 Euros of SO (ES6); from 100,000 to less than 250,000 Euros of SO (ES7); 250,000 Euros of SO and more (ES8 to ES14). It should be noted that, according to the EU regulation (CE) N$^{\underline{o}}$1242/2008, the two last categories correspond to only large farms. The large sized farms are divided in two classes to have at least 3 category of sizes. In the following, these farm categories are referred as medium, large and very large, respectively. As RICA being a rotating panel, farms which either enter or leave the sample in a given year cannot be considered as actual entries into or exits from the agricultural sector. Because of that constant population is assumed and only transitions between category of sizes as defined above will be investigated.

For estimation purposes, the sample is restricted to farms which were present in the database for at least two consecutive years to observe at least one transition. The corresponding unbalanced panel then comprised 13,325 farms out of the 15,841 farms in the original database (84.12%), leading to 89,229 (farm$\times$year) observations and 75,904 individual 1-year transitions (including staying in the same category of sizes) from 2000 to 2013. Table 1 shows that farms are more likely to remain in their initial category of sizes two consecutive years. More than 90% of farm remain in their initial category whatever the category of sizes considered. It should be noted that remaining in the initial category does not means that farms do not increase or decrease size, but the change is not sufficient to fall into a different category of sizes as defined in this study.

4.2. Explanatory variables

Several theoretical and empirical studies provide various factors that may play an important role on structural change in farming (see for example Goddard *et al.* (1993); Boehlje (1992); Harrington and Reinsel (1995)). These studies distinguish several categories among these factors. In the following, the selected causative factors of structural change are presented, focusing on some recent public support programs. The selection of the explanatory variables is based on the objective of the study and their availability in the database.

As the aim of this paper is to investigate the impacts of agricultural policies on farm structural change, subsidies received by farms from some public support programmes are used as explanatory variables of transition probabilities. Considering all farms all together, this study analyses the impacts of public support programmes mainly originate from the Common Agricultural Policy (CAP). Indeed, subsidies from the CAP are divided into two main components called 'pillars'. The aims of the support programmes from these two pillars are different from each order. Theoretically, no consensus has been found for the real effects of such public programs (Zimmermann and Heckelei, 2012). The impact of public supports may depends on how programs are designed for that commodity (Goddard *et al.* (1993)). The impacts of public supports from the two pillars of the CAP are thus separately investigated on the transition probabilities of farms.

Farm subsidies from the Pillar One are divided in coupled and decoupled subsidies. Despite the fact that decoupled subsidies were introduced in order to reduce the impacts on level of activity of farms, this kind of supports appear to increase the farm size on the long run. Indeed, decoupled subsidies may become coupled since the first ones are based on historical production. Therefore, farmers had better to increase their operate farm size if they expected more subsidies in the future. According to the literature, a positive impacts of subsidies from the Pillar One of the CAP is expected to favour farm growth, gathering coupled and decoupled subsidies.

Policy instruments from Pillar Two are devoted to promote rural development. Two types of subsidies are considered from Pillar Two according to potential impacts on farm size changes. The first component includes all subsidies allocated for the agri-environment and climate change programs. These kinds of supports may have a negative impacts on farm growth since agri-environment measures supposed to facilitate change in farming systems towards more resilient styles of production, better able to cope with future climate-related stress (Dwyer, 2013). The second component gather subsidies allocated for farm's investments, climatic damage and production diversification. Such kinds of support programs are more likely to favour farm size stagnation and growth by reducing risk. Gathering these two kinds of subsidies from the Pillar Two of the CAP, a positive impact is thus expected on farm size stagnation.

Several factors other than agricultural policy have been proved to affect structural change in farming (see Harrington and Reinsel (1995); Zimmermann *et al.* (2009)). In this study, only factors related to farm path dependency and economic environment of farms are considered. The reason for only considering these factors is because of data limitations. Indeed, proxies for other factors that may play an import role on farm size change, such as market condition, technical change, ect., are not available at individual farm level in the RICA database. Nevertheless, as factors that affect transition process of farms may relate to each other (Goddard *et al.*, 1993), proxies used may also capture impacts of some other causative factors of farm size change.

Following Zimmermann and Heckelei (2012), the total initial stocks is used as a proxy of path dependency. Initial stocks is supposed to negatively affect farm size declined since high initial stocks are assumed to result from former investment. Gross Operating Surplus (GOS) minus the total amount of subsidies received and the debt rate of farms were used to reflect the economic environment of farms. The GOS represents the financial capacity of a farm and as such is a very important indicator to obtain credit from a bank which can be used for new investments. The GOS could also relate to the self-financing capacity of farms. A positive effect of the Gross Operating Surplus is thus expected on the probability of farm to grow. Debt rate is also expected to have a positive impact on farm growth since credit generally enables firms to obtain necessary resources Goddard *et al.* (1993).

Finally, some farm and farmer characteristics such as age, managerial capacity of farmers and legal status, localisation, specialisation of farms are also used in the specification of transition probabilities. Farm and farmer characteristics may allow controlling for observed heterogeneity and are introduced in the model specification using dummy variables. Table 2 presents the description and summary statistics for the all chosen explanatory variables.

Before starting with the results, it should be noted that, given some potential sources of unobserved heterogeneity as mentioned in section 1, it supposed that impacts of explanatory variables, especially farm subsidies, may vary according to the types that a farm belong to.

4.3. Results

The mixed Markov model is applied to the RICA data described above. It is assumes that farms do not move from medium to very large size and vice versa since only few movements between these two category of sizes are observed (see Table 1). That is a common procedure in the agricultural economic literature when using Markov modelling approach (see Ben Arfa *et al.* (2015) for a recent example). The main results are presented and discussed in this section.

4.3.1. Type membership and transition probability matrices

According to the information criteria presented in section 3, a mixture of two types of farms seems to be the most appropriated data generating process. Two types are chosen as the optimal number specifically for two reasons: first,BIC and CAIC criteria indicate that two types is the optimal number of farm types; second, even AIC and AIC3 criteria indicate a higher optimal number of farm types, the results show that the improvement of these criteria is relatively small when specifying more than two types of farms (see Figure 1). This means that increasing the total number of homogeneous types in the population increase much more the total number of parameters to estimate than the representativeness of the data generating process. A mixture of two types of farms is thus preferable to represent farm size dynamics in France.

Table 3 reports the estimated type shares and the resulting transition probability matrices (TPMs) for the two farm types. As expected, the resulting TPMs are quite different from each other. The average posterior probabilities of belonging to a specific type indicate that a majority of farms tends to remain in their initial category of sizes indefinitely (at least during the entire period of observation). Indeed, about 68% of the sample consists of farms which predominantly stay in their initial category of sizes. Conversely, farms belonging to the second type, about 32% of the sample, are more likely to change category of sizes two consecutive years than farms in the first type. In the following, the first farm type is called 'almost stayers' and those in the second farm type 'movers'.

For the almost stayer farm type, the categories are almost all absorbing states (see Table 3.a). The probability to remain in the same category of sizes two consecutive years for those farms are close or over 0.99. This result means that these farms have about 99% of chance to remain in the same category of sizes during a long time period. The transition probability matrix of the 'movers' type is also strongly diagonal meaning that even farms which are likely to change category of sizes also have a high probability to remain in their initial category two consecutive years (table 3.b). However, the probability to remain in the same category of sizes for movers is around 0.85 meaning that farms in this type have about 15% more chance to change their category of sizes than those in the almost stayer type. In both farm types, large sized farms represent the largest category with about 43% of the total number of farms. These larger proportions of large sized farms in both types of farms could be explained by the fact that this category of sizes is the largest in the sample (see Table 1).

Summary statistics for various farm and farmer characteristics for both almost stayers and movers farm types are then computed in order to identify the profile of farms in each type. Table 5 show that farmers who are close retirement and farms specialized in crops or operated under individual legal status are more likely to be in the almost stayer type. Contrary, farms whose most of part is located in area without natural handicap (mountains, piedmont plains, etc.) and farms whose farmers have received at least a minimum agricultural training are more likely to behave as a mover. However, the probability of belonging to a specific type is not highly correlated to the farm and farmer characteristics considered in the model specification. Indeed, the statistics are quiet similar for all observed characteristics of farms and farmers in both types. The difference in the distribution of these characteristics between the two types of farms is very low. The difference in the proportion of farms with a specific observed characteristic in the two types is lower than 1% for almost all characteristics considered in the study. For example, the proportion of farmers over 55 years old in the stayer type and in the movers type are 17.80% and 17.50%, respectively. The descriptive statistics thus show that unobserved heterogeneity cannot be sufficiently controlled by some observed farms and/or farmer characteristics and accounting for both kinds of heterogeneity may therefore lead to more efficiently estimate the impacts of explanatory variables, including agricultural policies.

The transition probability matrix for the overall population of farms can be easily derived by summing the two types of TPMs weighted by their respective shares in the population. The resulting 1-year TPM for the overall population is reported in Table 4. The transition probabilities show that, on overall, farms are more likely to remain in their initial category two consecutive years which is a common features in agricultural economics (see for example (Hallberg, 1969; Stokes, 2006; Piet, 2011)). Indeed, the overall population TPM is strongly diagonal. The probability to remain in the same category of size two consecutive years is around 0.94. This high probability to remain in the same category of sizes is due to the high proportion of the stayer type in the population. As a consequence, considering an homogeneous population to describe farm size dynamics as well as to investigate the impact of some explanatory variables on transition probabilities may be not sufficiently informative. Analysis under a mixture approach may be more informative by separating the impacts of explanatory variables, including agricultural policies, on different farm patterns.

4.3.2. Impact of explanatory variables

The parameters of transition probabilities are estimated under the mixture Markov assumption. For each type of farms, a multinomial logit regression is estimated using the posterior probability of belonging to a specific type as a weighted factor as described in section 2.2. As mentioned in section 2.1, for each initial category of sizes the alternative to remain in the same category of sizes two consecutive years is used as the reference. The estimated coefficients for the odds ratios are reported in Table 6. As expected, the results show that the impacts of explanatory variables are different given the type of farms considered. Even the same sign is most often observed for some parameters, the coefficient values are generally different meaning that the impacts of explanatory variables depend on farm type membership. On overall, subsidies from the two pillars, initial stocks, total Gross Operating Surplus and debt rate of farms have a positive impact on the probability to grow and negatively affect farm size declined whatever the farm type considered. These results are consistent to the literature and confirm the expectations. Several studies showed that subsidies from some support programs of the Common Agricultural Policy (CAP) are likely to favour farm growth (see Ben Arfa *et al.* (2015) for a recent example). Zimmermann and Heckelei (2012) found that stocking density positively affect farm growth. Contrary to the later authors, the results also support that initial stocks decrease the probability of farms to decline. Furthermore, it should be noted that subsidies from the Pillar Two are more likely to increase farm probability to grow only for medium sized farms. These kinds of subsidies generally decrease the probability to decline. This result confirms the expectation that subsidies from the Pillar Two of the CAP are less likely to encourage farm growth than subsidies from the Pillar One.

The results also show that farms are more likely to decrease category of sizes when farmers are over 55 years old, in both types of farms. Individual legal status as well as specialisation in crop productions have the same impacts on farm growth. These results can be explained by the fact that: first, farmers may be less motivated to increase their capacity of production when they are close of retirement because they should be more interested to the farm succession (Potter and Lobbley, 1992) or by the fact that younger farmers are more likely to seek to increase agricultural activity, as they would be less financially secure than their older counterparts (Howley *et al.*, 2014); second, farms involving under individual legal status may face much more economical constraints (capital, access to credit, etc.) than corporate farms which may constrain new investments or such farms may just have less financial motives than corporate farms (Boehlje, 1992); third, it may be more difficult for farms specialized in crop productions to increase their operated farm size over time because of the regulation of the land market, specially in France, since increasing the production capacity for crop farms may generally need to increase the total land used more than for livestock production systems. Conversely to the previous farm and farmer characteristics, the probability to grow increase if the most part of the farm is located in an area without natural handicap or if farmers have received a minimum agricultural trainings. The later result confirm the positive impact of the managerial capacity on farm size growth (Boehlje, 1992; Goddard *et al.*, 1993).

Tables 8, 9 and 10 respectively report the probability elasticities for farms to grow, decline and remain in the same category of sizes two consecutive years. The results show that the impact of subsidies both from the Pillar One and Pillar Two of the CAP on farm probability to grow or decline is higher for farms belonging to the almost stayer type than for farms in the movers one. For example, a 1% increase of subsidies from the Pillar One (in 1000 Euros) will increase the probability of farms in the almost stayer type to move from medium sized to large about 0.57 but only 0.17 for the movers type. The impact of coupled and decoupled subsidies is thus about 70% higher for farms in the almost stayers than for the movers ones. This result may be explained by the fact that the almost stayer type could gather farms who have some liquidity constraints and aids from government programs may have an income multiplier effect for those farms (Latruffe *et al.*, 2010).

Nevertheless, the results also show that for the almost stayer type of farms the subsidies from the Pillar One of the CAP only have a positive impact on growth of medium sized farms. Indeed, large and very large sized farms in this type are more likely to remain in their category of sizes when increased their amount of subsidies. It could be the fact that farms in the almost stayer type may be less motivated to increase their operated farm size at a certain level for some of the reasons mentioned in section 1. In particular, they may have less financial motives and therefore may have a lower optimal size than farms in the movers type. As the subsidies received will increase the total profit, the probability for these farms to become very large may decrease with the amount of subsidies received.

Structure elasticities are then computed at the mean effects of subsidies from both pillars. Since it is impossible to take into account entry and exit in farming and constant population were assumed, the structure elasticities are computed on farm proportion by category of sizes. The structure elasticities thus represent the variation of farm proportion in a specific category of sizes that will occur by 1% change of the total amount of subsidies received by farms. Figure 2 and 3 present the resulting structure elasticities for both the homogeneous and the mixture models. The figures show that the resulting structure elasticities from the homogeneous model and the mixture one are quite different. Overall, the homogeneous model tends to overestimate the impacts of subsidies from both Pillar one and Pillar Two of the CAP. For example, the homogeneous model predicts that a 1% increase of subsidies from the Pillar One will decrease the proportion of medium sized farms by about 2.72% while the mixture model predicts a decrease only about 1.82% (see Figure 2). Likewise, the homogeneous model predicts 0.67% increase of the proportion of very large sized farm if subsidies from the Pillar Two increase by 1% (see Figure 3) while this proportion is only about 0.41% for the mixture model. These results confirmed that ignoring unobserved heterogeneity in modelling farm size dynamics can lead to different level of agricultural policy impacts on structural change and thus to misleading results.

5. Concluding remarks

This paper provides a new approach for modelling structural change in agriculture. Using individual level data, a mixture of homogeneous farm types is considered in order to incorporate unobserved heterogeneity in farm transition process. A discrete choice modelling approach is used to describe farm size choices by farmers as decision makers. A multinomial logit specification of the transitions probabilities were applied and the expectation-maximization algorithm allowed estimating parameters of the mixed Markov chain model. Transition probability elasticities as well as structure elasticities are derived to analyse the impacts of some exogenous variables on farm size change in the French farming sector, specially focusing on the impact of some agricultural policies.

Using a sample of farms from the 'Réseau d'Information Compatable Agricole' (RICA) from 2000 to 2013, the results showed that French farms can be divided in two types: 'almost stayer' who are more likely to remain in their initial category of sizes and 'movers' who change category of sizes more frequently. The results also showed that the French farm population consists of a higher proportion of farms that behave like almost stayer leading to a strongly diagonal transition probability matrix for the overall population. Descriptive statistics showed that the probability to belong to the almost stayer type or the movers one are not very correlated with the observed farm and/or farmer characteristics meaning that unobserved heterogeneity cannot be fully controlled by observed one.

The results also showed that the impacts of farm subsidies from Pillar One and Pillar Two of the Common Agricultural Policy (CAP), depends on the type that farms belong to. Aggregated at the population level, structure elasticities showed that mixture model leads to different results and overall the homogeneous model tends to overestimate the impacts of subsidies both from Pillar One and Pillar Two on farm size change over time. These results are relevant for policy assessment since they confirmed that ignoring potential unobserved heterogeneity in farmer behaviour may lead to incorrect parameters and therefore to misleading results.

This study has some limitations that may motivate further research. In the current paper, the estimations were performed under constant population assumption since it is impossible to take into account entries and exits in the French farming sector because of data limitations. Accounting for entry and exits could be an obvious way to analyse the impact of agricultural policy on total number of farms by category of sizes. The results could be also improved considering other explanatory variables than those used in this application since it has been proved that several other factors may play an important role on structural change in farming.

References

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.

Baudry, J.-P. and Celeux, G. (2015). Em for mixtures-initialization requires special care .

Ben Arfa, N., Daniel, K., Jacquet, F. and Karantininis, K. (2015). Agricultural policies and structural change in French dairy farms: A nonstationary Markov model. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 63: 19–42.

Boehlje, M. (1992). Alternative models of structural change in agriculture and related industries. *Agribusiness (1986-1998)* 8: 219.

Bostwick, D. (1962). *Yield probabilities as a Markov process*. Marketing Economics Division, Economic Research Service, US Department of Agriculture.

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

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.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 39: 1–38.

Dias, J. G. and Willekens, F. (2005). Model-based clustering of sequential data with an application to contraceptive use dynamics. *Mathematical population studies* 12: 135–157.

Dwyer, J. (2013). Transformation for sustainable agriculture: what role for the second pillar of cap? *Bio-based and Applied Economics* 2: 29–47.

Freshwater, D. and Reimer, B. (1995). Socio-economic policies as causal forces for the structure of agriculture. *Canadian journal of agricultural economics* .

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. H. (2006). *Econometric analysis*. Pearson Education India, 6th ed.

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

Hallberg, M. C. (1969). Projecting the size distribution of agricultural firms. an application of a markov process with non-stationary transition probabilities. *American Journal of Agricultural Economics* 51: 289–302.

Harrington, D. H. and Reinsel, R. D. (1995). synthesis of forces driving structural change. *Canadian Journal of Agricultural Economics* .

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

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* .

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.

Huettel, S. and Jongeneel, R. (2011). How has the EU milk quota affected patterns of herd-size change? *European Review of Agricultural Economics* 38: 497–527.

Krenz, R. D. (1964). Projection of farm numbers for North Dakota with Markov chains. *Agr. Econ. Res.* 16: 77–83.

Latruffe, L., Davidova, S., Douarin, E. and Gorton, M. (2010). Farm expansion in lithuania after accession to the eu: The role of cap payments in alleviating potential credit constraints. *Europe-Asia Studies* 62: 351–365.

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.

Pacifico, D. and Yoo, H. I. (2012). lclogit: a stata module for estimating latent class conditional logit models via the expectation-maximization algorithm. *UNSW Australian School of Business Research Paper* .

Padberg, D. I. (1962). The use of Markov processes in measuring changes in market structure. *J. Farm Econ.* 44: 189–199.

Piet, L. (2011). Assessing structural change in agriculture with a parametric Markov chain model. Illustrative applications to EU-15 and the USA. Paper presented at the XIIIth Congress of the European Association of Agricultural Economist, Zurich (Switzerland).

Potter, C. and Lobbley, M. (1992). Ageing and succession on family farms: The impact on decision-making and land use. *Sociologia Ruralis* 32: 317–334.

Rahelizatovo, N. C. and Gillespie, J. M. (1999). Dairy farm size, entry, and declining production region. *J. Agr. Appl. Econ.* 31: 333–347.

Saint-Cyr, L. D. and Piet, L. (2014). Movers and Stayers in the Farming Sector: Another Look at Heterogeneity in Structural Change. In *2014 International Congress, August 26-29, 2014, Ljubljana, Slovenia*, 183068. European Association of Agricultural Economists.

Stokes, J. R. (2006). Entry, exit, and structural change in pennsylvania's dairy sector. *Agr. Resource Econ. Rev.* 35: 357.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

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* .

Vermunt, J. (2010). Longitudinal research using mixture models. In Montfort, K. van, Oud, J. H. and Satorra, A. (eds), *Longitudinal Research with Latent Variables*. Springer Berlin Heidelberg, 119–152.

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.

Zepeda, L. (1995). Asymmetry and nonstationarity in the farm size distribution of Wisconsin milk producers: An aggregate analysis. *American Journal of Agricultural Economics* 77: 837–852.

Zimmermann, A. and Heckelei, T. (2012). Structural change of European dairy farms: A cross-regional analysis. *Journal of Agricultural Economics* 63: 576–603.

Zimmermann, A., Heckelei, T. and Dominguez, I. P. (2009). Modelling farm structural change for integrated ex-ante assessment: Review of methods and determinants. *Environmental Science and Policy* 12: 601–618.

Table 1. Farm 1-year interval transitions across category of sizes, 2000-2013

		ES class			Total transitions
		25-100	100-250	≥ 250	
ES class	25-100	26,034	1,310	27	27,371
	100-250	1,257	32,425	1,055	34,737
	≥ 250	31	840	12,925	13,796

Note: ES in 1,000 Euros of standard output (SO).

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 2. Descriptive statistics and definition of explanatory variables (n=75,904)

Variable	Mean	Std. Dev.	Min.	Max.
pillar1	30.271	29.374	0.000	315.404
(Subsidies from Pillar One)				
pillar2	3.615	7.150	0.000	96.3
(Subsidies from Pillar Two)				
in_ stocks	97.563	158.808	0.000	3,487.727
(Total initial stocks)				
GOS/utans	26.937	43.352	-1,052.018	1,067.246
(GOS per non-salaried UTA)				
debt_ rate	39.678	26.679	0.000	545.300
(Total debt/liabilities)				
close_ retirement	0.177	0.381	0.000	1.000
(over 55 years old[a])				
crops	0.451	0.498	0.000	1.000
(Specialized in crops)				
individual	0.513	0.500	0.000	1.000
(Operate under individual legal status)				
well_ located	0.599	0.490	0.000	1.000
(Most part in area without natural handicap)				
agri_ skills	0.936	0.246	0.000	1.000
(Has agricultural trainings)				

pillar1, *pillar2*, *in_ stock* and *GOS/utans* in 1,000 Euros;
[a]In order to take into account farmers' anticipation, 55 years old is used instead of 62 which is the legal age of retirement in France.

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 3. Estimated farm type shares and 1-year transition probability matrices (TPMs)

		Shares	ES class		
		68.15%	25-100	100-250	≥ 250
ES class	25-100	40.10%	0.992(0.014)	0.008(0.014)	0.000(.)
	100-250	42.92%	0.007(0.009)	0.986(0.012)	0.007(0.007)
	≥ 250	16.98%	0.000(.)	0.011(0.019)	0.989(0.019)

a) Almost Stayers TPM

		Shares	ES class		
		31.85%	25-100	100-250	≥ 250
ES class	25-100	36.90%	0.858(0.113)	0.142(0.113)	0.000(.)
	100-250	43.12%	0.091(0.076)	0.828(0.080)	0.081(0.074)
	≥ 250	19.98%	0.000(.)	0.151(0.100)	0.849(0.100)

a) Movers TPM

Note: ES in 1,000 Euros of standard output (SO); standard deviations in parenthesis

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 4. Overall population 1-year transition probability matrix.

		ES class		
		25-100	100-250	≥ 250
ES class	25-100	0.951(0.045)	0.049(0.045)	0.000(.)
	100-250	0.034(0.029)	0.935(0.027)	0.031(0.025)
	≥ 250	0.000(.)	0.055(0.038)	0.945(0.038)

Note: ES in 1,000 Euros of standard output (SO); standard deviations in parenthesis

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 5. Descriptive statistics for observed farm characteristics by type membership.

Variable	Almost Stayers		Movers	
close_retirement	0.822	(0.382)	0.825	(0.380)
crops	0.453	(0.498)	0.447	(0.497)
individual	0.516	(0.500)	0.506	(0.500)
well_located	0.599	(0.490)	0.600	(0.490)
agri_skills	0.933	(0.249)	0.940	(0.238)

Standard deviations in parenthesis

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 6. Estimated parameters of farm probabilities to grow two consecutive years for the homogeneous and the mixed Markov models.

Odds ratios	Variables	Homogeneous		Almost Stayers		Movers	
p12/p11	*intercept*	-3.912***	(0.208)	-6.804***	(0.385)	-2.998***	(0.193)
	pillar1	0.274***	(0.044)	0.417***	(0.071)	0.141***	(0.042)
	pillar2	-0.119	(0.071)	0.229**	(0.080)	-0.206**	(0.070)
	in_stocks	0.003 ***	(0.001)	0.004***	(0.001)	0.013***	(0.001)
	GOS/utans	0.010***	(0.002)	0.005*	(0.002)	0.011***	(0.002)
	debt_rate	0.010***	(0.001)	0.010***	(0.001)	0.015***	(0.001)
	close_retirement	0.307**	(0.100)	1.046***	(0.214)	0.150	(0.096)
	crops	-0.387***	(0.081)	-0.254	(0.137)	-0.363***	(0.072)
	individual	-0.783***	(0.074)	-1.412***	(0.115)	-0.714***	(0.068)
	well_located	0.283***	(0.083)	0.641***	(0.134)	0.249***	(0.070)
	agri_skills	0.285*	(0.146)	0.362	(0.264)	0.314*	(0.137)
p23/22	*intercept*	-4.782***	(0.242)	-4.697***	(0.342)	-4.400***	(0.263)
	pillar1	0.169***	(0.044)	-0.536***	(0.057)	0.311***	(0.039)
	pillar2	0.055	(0.053)	-0.049	(0.089)	0.024	(0.058)
	in_stocks	0.003***	(0.000)	0.001	(0.001)	0.005***	(0.000)
	GOS/utans	0.006***	(0.001)	0.003**	(0.001)	0.008***	(0.001)
	debt_rate	0.012***	(0.001)	0.009***	(0.002)	0.014***	(0.001)
	close_retirement	0.096	(0.096)	-0.030	(0.150)	0.164	(0.099)
	crops	-0.436***	(0.081)	0.220	(0.122)	-0.750	(0.083)
	individual	-0.520***	(0.082)	-0.040	(0.119)	-0.716***	(0.083)
	well_located	0.200*	(0.087)	0.180	(0.146)	0.167***	(0.083)
	agri_skills	0.222	(0.184)	-0.201	(0.240)	0.384***	(0.210)

Note: Standard errors in parenthesis; ***, ** and * are significant at 0.1%, 1% and 5%, respectively.

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 7. Estimated parameters of farm probabilities to decline two consecutive years for the homogeneous and the mixed Markov models.

Odds ratios	Variables	Homogeneous		Almost Stayers		Movers	
p21/p22	*intercept*	-1.471***	(0.195)	-4.509***	(0.326)	0.228	(0.185)
	pillar1	-0.224***	(0.037)	-0.377***	(0.066)	-0.161***	(0.035)
	pillar2	-0.053	(0.040)	0.045	(0.063)	-0.083	(0.043)
	in_stocks	-0.009***	(0.001)	0.002*	(0.001)	-0.013***	(0.001)
	GOS/utans	-0.014***	(0.002)	-0.013***	(0.002)	-0.014***	(0.002)
	debt_rate	-0.005***	(0.002)	0.008***	(0.002)	-0.009***	(0.001)
	close_retirement	-0.327***	(0.081)	-1.068***	(0.127)	-0.163	(0.085)
	crops	-0.158*	(0.075)	-0.379**	(0.122)	-0.104	(0.071)
	individual	0.802***	(0.075)	0.761	(0.122)	0 .675***	(0.070)
	well_located	-0.485***	(0.075)	0.013	(0.130)	-0.539***	(0.073)
	agri_skills	-0.105	(0.149)	0.536*	(0.249)	-0.526***	(0.141)
p32/p33	*intercept*	-1.066***	(0.252)	-2.434***	(0.375)	0.469*	(0.244)
	pillar1	-0.060	(0.038)	-0.355***	(0.067)	-0.122**	(0.040)
	pillar2	-0.201**	(0.067)	-0.139	(0.097)	-0.203**	(0.071)
	in_stocks	-0.004***	(0.001)	0.001	(0.001)	-0.006***	(0.001)
	GOS/utans	-0.007***	(0.001)	-0.01***	(0.002)	-0.007***	(0.001)
	debt_rate	-0.006***	(0.002)	-0.003	(0.003)	-0.009***	(0.002)
	close_retirement	-0.131	(0.099)	-1.111***	(0.159)	-0.073	(0.105)
	crops	-0.266**	(0.098)	-1.170***	(0.190)	-0.037	(0.084)
	individual	0.431***	(0.099)	1.914***	(0.154)	0.147	(0.091)
	well_located	-0.261*	(0.102)	0.055	(0.233)	-0.212*	(0.096)
	agri_skills	0.018	(0.188)	-0.659**	(0.257)	-0.013	(0.171)

Note: Standard errors in parenthesis; ***, ** and * are significant at 0.1%, 1% and 5%, respectively.

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 8. Yearly probability elasticities of farm growth both for homogeneous and mixed Markov chain models.

Probability	Variables	Homogeneous		Almost Stayers		Movers	
Medium/Large (p12)	*pillar1*	0.363	(0.058)	0.574	(0.098)	0.171	(0.051)
	pillar2	-0.058	(0.035)	0.118	(0.041)	-0.090	(0.031)
	i*in_ stocks*	0.111	(0.031)	0.155	(0.045)	0.381	(0.026)
	GOS/utans	0.103	(0.021)	0.056	(0.024)	0.103	(0.020)
	debt_ rate	0.298	(0.027)	0.323	(0.043)	0.393	(0.026)
	close_ retirement	0.241	(0.078)	0.855	(0.175)	0.105	(0.068)
	crops	-0.149	(0.032)	-0.101	(0.055)	-0.128	(0.026)
	individual	-0.594	(0.057)	-1.115	(0.091)	-0.491	(0.048)
	well_ located	0.118	(0.034)	0.277	(0.058)	0.093	(0.026)
	agri_ skills	0.248	(0.127)	0.329	(0.241)	0.246	(0.106)
Large/V. large (p23)	*pillar1*	0.304	(0.075)	-0.955	(0.102)	0.522	(0.062)
	pillar2	0.020	(0.019)	-0.018	(0.031)	0.012	(0.020)
	in_ stocks	0.231	(0.028)	0.085	(0.056)	0.421	(0.028)
	GOS/utans	0.155	(0.028)	0.095	(0.034)	0.218	(0.024)
	debt_ rate	0.480	(0.049)	0.372	(0.069)	0.558	(0.055)
	close_ retirement	0.086	(0.077)	-0.020	(0.123)	0.137	(0.076)
	crops	-0.197	(0.037)	0.103	(0.057)	-0.326	(0.037)
	individual	-0.227	(0.033)	-0.019	(0.047)	-0.338	(0.036)
	well_ located	0.136	(0.055)	0.118	(0.095)	0.125	(0.049)
	agri_ skills	0.208	(0.169)	-0.192	(0.225)	0.384	(0.186)

Note: Standard errors in parenthesis

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 9. Yearly probability elasticities of farm declined both for homogeneous and mixed Markov chain models.

Probability	Variables	Homogeneous		Stayers type		Movers type	
Large/Medium (p21)	*pillar1*	-0.397	(0.065)	-0.669	(0.118)	-0.312	(0.058)
	pillar2	-0.019	(0.014)	0.016	(0.022)	-0.029	(0.015)
	in_stocks	-0.751	(0.107)	0.148	(0.062)	-1.070	(0.079)
	GOS/utans	-0.366	(0.042)	-0.344	(0.052)	-0.368	(0.055)
	debt_rate	-0.224	(0.062)	0.342	(0.092)	-0.386	(0.057)
	close_retirement	-0.264	(0.065)	-0.878	(0.104)	-0.133	(0.064)
	crops	-0.067	(0.034)	-0.178	(0.057)	-0.024	(0.031)
	individual	0.315	(0.028)	0.300	(0.048)	0.264	(0.026)
	well_located	-0.312	(0.048)	0.008	(0.085)	-0.33	(0.045)
	agri_skills	-0.102	(0.136)	0.504	(0.233)	-0.483	(0.124)
V. Large/Large (p32)	*pillar1*	-0.106	(0.067)	-0.643	(0.121)	-0.206	(0.069)
	pillar2	-0.073	(0.024)	-0.054	(0.038)	-0.063	(0.023)
	in_stocks	-0.794	(0.150)	0.209	(0.114)	-1.171	(0.110)
	GOS/utans	-0.379	(0.055)	-0.530	(0.097)	-0.325	(0.051)
	debt_rate	-0.269	(0.080)	-0.167	(0.168)	-0.383	(0.071)
	close_retirement	-0.100	(0.076)	-0.886	(0.127)	-0.051	(0.073)
	crops	-0.127	(0.047)	-0.601	(0.098)	-0.015	(0.035)
	individual	0.089	(0.020)	0.421	(0.033)	0.026	(0.016)
	well_located	-0.191	(0.075)	0.043	(0.182)	-0.138	(0.063)
	agri_skills	0.016	(0.167)	-0.612	(0.239)	-0.011	(0.138)

Note: Standard errors in parenthesis

Source: Agreste, RICA France 2000-2013 – authors' calculations

Table 10. Yearly probability elasticities of farms to remain in the same category of sizes both for homogeneous and mixed Markov chain models.

Probability	Variables	Homogeneous		Stayers type		Movers type	
Medium/Medium (p11)	*pillar1*	-0.021	(0.004)	-0.006	(0.001)	-0.031	(0.010)
	pillar2	0.003	(0.001)	-0.001	(0.001)	0.012	(0.004)
	in_ stocks	-0.008	(0.003)	-0.003	(0.001)	-0.088	(0.008)
	GOS/utans	-0.008	(0.002)	-0.001	(0.000)	-0.025	(0.006)
	debt_ rate	-0.020	(0.002)	-0.004	(0.001)	-0.087	(0.007)
	close_ retirement	-0.013	(0.004)	-0.008	(0.002)	-0.018	(0.012)
	crops	0.007	(0.001)	0.001	(0.000)	0.020	(0.003)
	individual	0.023	(0.002)	0.005	(0.000)	0.064	(0.005)
	well_ located	-0.007	(0.002)	-0.003	(0.001)	-0.018	(0.005)
	agri_ skills	-0.013	(0.007)	-0.003	(0.002)	-0.042	(0.018)
Large/Large (p22)	*pillar1*	0.003	(0.003)	0.008	(0.001)	-0.027	(0.009)
	pillar2	0.000	(0.001)	0.000	(0.000)	0.003	(0.002)
	in_ stocks	0.008	(0.002)	-.002	(0.001)	0.018	(0.006)
	GOS/utans	0.001	(0.002)	0.000	(0.000)	-0.002	(0.004)
	debt_ rate	-0.010	(0.003)	-0.006	(0.001)	-0.021	(0.008)
	close_ retirement	0.007	(0.003)	0.005	(0.001)	0002	(0.009)
	crops	0.008	(0.001)	0.000	(0.001)	0.025	(0.003)
	individual	-0.014	(0.002)	-0.003	(0.001)	-0.028	(0.005)
	well_ located	0.005	(0.002)	-0.001	(0.001)	0.018	(0.005)
	agri_ skills	-0.003	(0.008)	-0.002	(0.002)	0.018	(0.019)
V.Large/V.Large (p33)	*pillar1*	0.007	(0.004)	0.004	(0.001)	0.038	(0.012)
	pillar2	0.004	(0.001)	0.000	(0.000)	0.010	(0.003)
	in_ stocks	0.028	(0.004)	-0.002	(0.002)	0.104	(0.007)
	GOS/utans	0.015	(0.002)	0.004	(0.000)	0.035	(0.004)
	debt_ rate	0.017	(0.005)	0.002	(0.002)	0.069	(0.012)
	close_ retirement	0.006	(0.005)	0.007	(0.001)	0.009	(0.013)
	crops	0.006	(0.002)	0.006	(0.001)	0.002	(0.005)
	individual	-0.008	(0.002)	-0.013	(0.002)	-0.006	(0.004)
	well_ located	0.011	(0.004)	0.000	(0.002)	0.023	(0.010)
	agri_ skills	-0.001	(0.011)	0.006	(0.002)	0.002	(0.025)

Note: Standard errors in parenthesis

Source: Agreste, RICA France 2000-2013 – authors' calculations

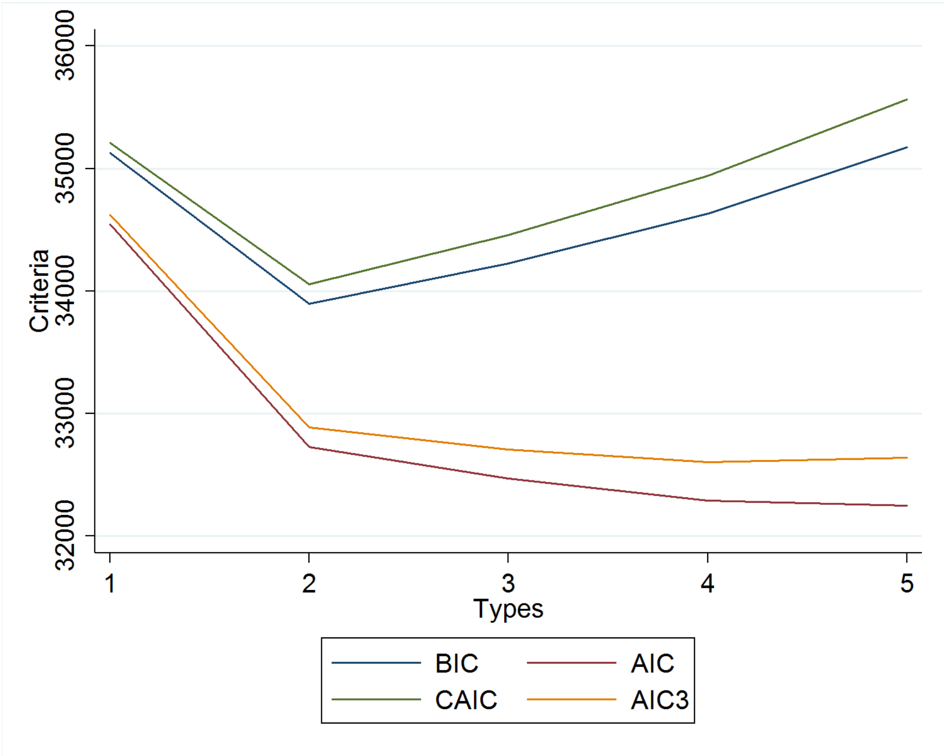

Figure 1. Comparison of model-fit statistics for different numbers of types.
Source: Agreste, RICA France 2000-2013 – authors' calculations

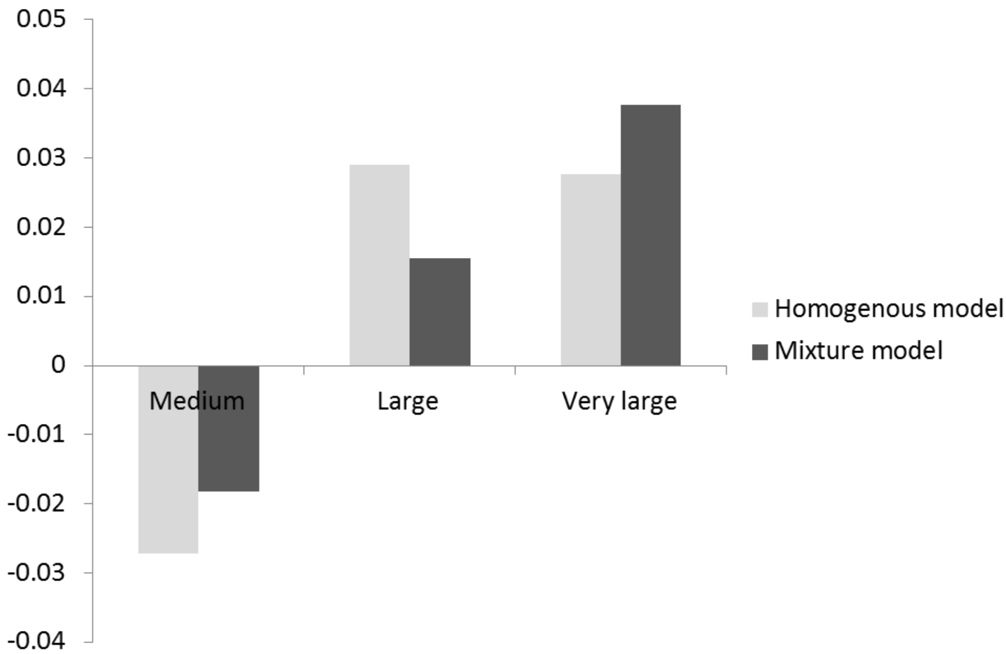

Figure 2. Yearly structure elasticities of farm subsidies from Pillar One of the CAP both for homogeneous and mixed Markov chain models.

Source: Agreste, RICA France 2000-2013 – authors' calculations

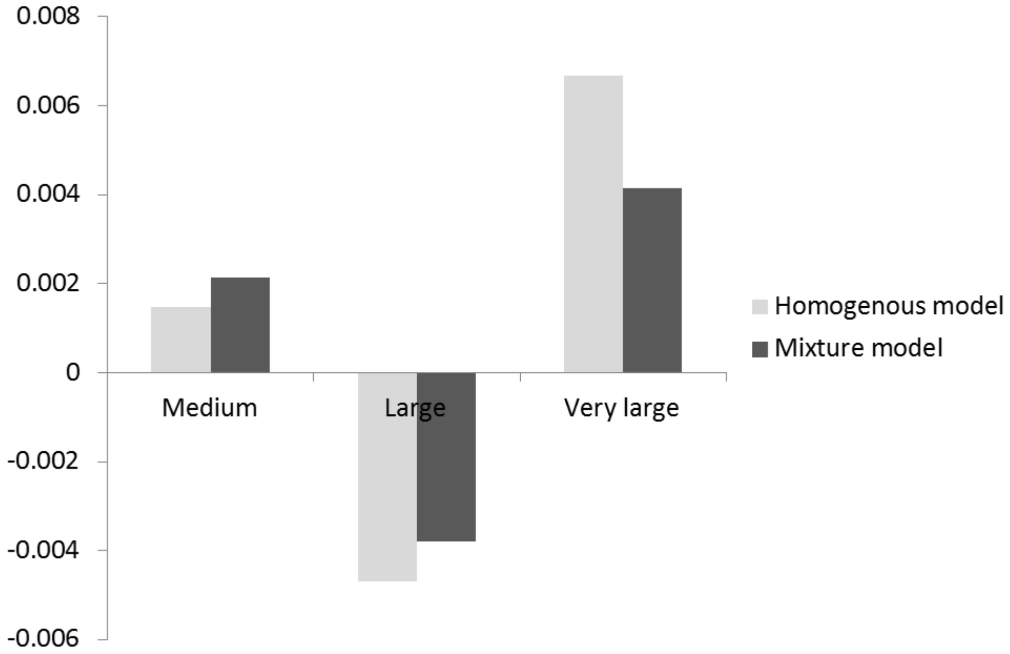

Figure 3. Yearly structure elasticities of farm subsidies from Pillar Two of the CAP both for homogeneous and mixed Markov chain models.

Source: Agreste, RICA France 2000-2013 – authors' calculations