Working Papers SMART - LERECO

Farm heterogeneity and agricultural policy impacts on size dynamics: evidence from France

Legrand D. F. SAINT-CYR

Working Paper SMART – LERECO N°17-04

June 2017

UMR INRA-Agrocampus Ouest **SMART - LERECO**

(Laboratoires d'Etudes et de Recherches en Economie sur les Structures et Marchés Agricoles, Ressources et Territoires)

Les Working Papers SMART-LERECO ont pour vocation de diffuser les recherches conduites au sein des unités SMART et LERECO dans une forme préliminaire permettant la discussion et avant publication définitive. Selon les cas, il s'agit de travaux qui ont été acceptés ou ont déjà fait l'objet d'une présentation lors d'une conférence scientifique nationale ou internationale, qui ont été soumis pour publication dans une revue académique à comité de lecture, ou encore qui constituent un chapitre d'ouvrage académique. Bien que non revus par les pairs, chaque working paper a fait l'objet d'une relecture interne par un des scientifiques de SMART ou du LERECO et par l'un des deux éditeurs de la série. Les Working Papers SMART-LERECO n'engagent cependant que leurs auteurs.

The SMART-LERECO Working Papers are meant to promote discussion by disseminating the research of the SMART and LERECO members in a preliminary form and before their final publication. They may be papers which have been accepted or already presented in a national or international scientific conference, articles which have been submitted to a peer-reviewed academic journal, or chapters of an academic book. While not peer-reviewed, each of them has been read over by one of the scientists of SMART or LERECO and by one of the two editors of the series. However, the views expressed in the SMART-LERECO Working Papers are solely those of their authors.

Farm heterogeneity and agricultural policy impacts on size dynamics: evidence from France

Legrand D. F. SAINT-CYR

SMART, AGROCAMPUS OUEST, INRA, 35000, Rennes, France

Corresponding author

Legrand D. F. Saint-Cyr
UMR SMART-LERECO
4 allée Adolphe Bobierre, CS 61103
35011 Rennes cedex, France
Email : legrand.saint-cyr@inra.fr
Téléphone/phone: +33 (0)2 23 48 54 18
Fax: +33 (0)2 23 48 53 80

Farm heterogeneity and agricultural policy impacts on size dynamics: evidence from France

Abstract

This article investigates the impact of agricultural policies on structural change in farming. Since not all farmers may behave alike, a non-stationary mixed-Markov chain modeling (M-MCM) approach is applied to capture unobserved heterogeneity in the transition process of farms. A multinomial logit specification is used for transition probabilities and the parameters are estimated by the maximum likelihood method and the Expectation-Maximization (EM) algorithm. An empirical application to an unbalanced panel dataset from 2000 to 2013 shows that French farming mainly consists of a mixture of two farm types characterized by specific transition processes. The main finding is that the impact of farm subsidies from both pillars of the Common Agricultural Policy (CAP) highly depends on the farm type. A comparison between the non-stationary M-MCM and a homogeneous non-stationary MCM shows that the latter model leads to either overestimation or underestimation of the impact of agricultural policy on change in farm size. This suggests that more attention should be paid to both observed and unobserved farm heterogeneity in assessing the impact of agricultural policy on structural change in farming.

Keywords: agricultural policy, EM algorithm, farm structural change, mixed-Markov chain model, multinomial logit, unobserved heterogeneity

JEL classification: Q12, Q18, C38, C51

Hétérogénéité des exploitations et impact des politiques agricoles sur l'évolution de la taille: étude de cas de la France

Résumé

Cet article analyse l'impact des politiques agricoles sur le changement structurel en agriculture. Etant donné que les agriculteurs ne se comportent pas tous de la même façon, une approche de mélange de modèles de chaînes de Markov (M-MCM) non-stationnaire est appliquée pour capturer l'hétérogénéité inobservée dans le processus de transition des exploitations agricoles. Un logit multinomial est utilisé pour spécifier les probabilités de transition et les paramètres sont estimés par du maximum de vraisemblance et l'algorithme d'Espérance-Maximisation (EM). Une application empirique sur un panel non cylindré de 2000 à 2013 montre que le secteur agricole en France est constitué principalement d'un mélange de deux types d'exploitations caractérisés par des processus de transition différents. Le principal résultat est que l'impact des subventions provenant des deux piliers de la Politique Agricole Commune (PAC) dépend fortement du type auquel appartiennent les exploitations. Une comparaison de ces résultats avec ceux obtenus avec un MCM homogène non-stationnaire montre que ce dernier conduit soit à surestimer, soit à sous-estimer l'impact de certaines subventions de la PAC sur l'évolution de la taille des exploitations. Cela suggère qu'il faudrait prêter plus d'attention à l'hétérogénéité, à la fois observée et inobservée, des exploitations agricoles dans l'évaluation de l'impact des politiques agricoles sur le changement structurel en agriculture.

Mots-clés : politiques agricoles, algorithme EM, changement structurel, modèles de mélange de chaînes de Markov, logit multinomial, hétérogénéité inobservée

Classification JEL : Q12, Q18, C38, C51

Farm heterogeneity and agricultural policy impacts on size dynamics: evidence from France

1 Introduction

The farming sector has faced important structural change over recent decades. In particular, the number of farms has been decreasing sharply and their average size measured in hectare has been increasing continually in most developed countries, implying changes in farm-size distribution. Such changes may have important consequences for equity among farmers, productivity and the efficiency of farming (Weiss, 1999; Breustedt and Glauben, 2007). Therefore, structural change has been the subject of considerable interest to agricultural economists and policy makers. Many theoretical studies pointed out the potential impacts of agricultural policies on farm-size changes, but these impacts remain ambiguous (Goddard *et al.*, 1993; Harrington and Reinsel, 1995). Since then, several studies have empirically investigated the impact of some recent policy programs on structural change in various farming contexts.

It has become common in agricultural economics to study structural change in farming and the impact of time-varying variables including agricultural policies by using the so-called Markov chain model (MCM). Focusing on farm number and size evolution, the Markovian framework has proven to be a convenient modeling approach to represent the transition process of farms across categories (Bostwick, 1962; Padberg, 1962; Krenz, 1964). Basically, this model states that the farm size in a given time period is the result of a probabilistic process where future farm size depends only on its size in the immediately previous period since, in general, a first order process is assumed. Stokes (2006) showed that a Markovian transition process may derive from a structural model of inter-temporal profit maximization, giving theoretical grounds for the use of the MCM.

This article aims at investigating the impacts of farm subsidies from the Common Agricultural Policy (CAP) on farm structural change, especially in France. It adds to the existing literature in two ways. First, a mixture modeling framework is applied to take into account potential unobserved farm heterogeneity in the analysis of structural change in farming. Second, transition probabilities are specified at the individual farm level within a discrete choice modeling approach in order to analyze the transition process at the farm level.

In agricultural economics, the heterogeneity issue has so far mostly been left in the background of farm structural change analysis. Two reasons may explain why there have been so few studies that explicitly account for farm heterogeneity. First, most previous studies focused on specific farm types. For example, several studies analyzed structural change in dairy farming, especially in some regions of the United States (Chavas and Magand, 1988; Zepeda, 1995a; Stokes, 2006), in Ireland (Keane, 1991; Gaffney, 1997; McInerney and Garvey, 2004), and in France (Ben Arfa

et al., 2015). Other analyses focused on other types of farming such as pig farms (Judge and Swanson, 1962; Disney *et al.*, 1988; Gillespie and Fulton, 2001; Karantininis, 2002), and cash grain farms (Bostwick, 1962; Garcia *et al.*, 1987). These authors may therefore have assumed that such groups are homogeneous enough to justify discarding heterogeneity issues. Second, studies that investigate the potential drivers of structural change in farming emphasize the role of economic factors such as market or policy variables. Researchers have therefore plainly left heterogeneity issues in the background, considering farm heterogeneity to be controlled by individual characteristics included in the model specification. Furthermore, previous studies of farm structural change often faced data limitations. Indeed, structural change in farming has generally been investigated using aggregate data (i.e., cross-sectional observations of the farm distribution across a finite number of size categories), because such data are most often available to researchers. Thus, it could be complicated, or even impossible, to control for farm heterogeneity using such data.

However, accounting for heterogeneity may be crucial to fully understand structural change in farming since this results from individual farmers' decisions (Freshwater and Reimer, 1995; Bollman *et al.*, 1995; Jackson-Smith, 1999). Farm heterogeneity may originate from several sources. One of the most important sources is farmers' motivation. While the standard economic theory of production states that farmers are normally expected to maximize their total profit from farming activities, it has been shown that not all farmers actually prioritize to profit maximization (Maybery *et al.*, 2005; Mzoughi, 2011; Howley *et al.*, 2014). This is the case for environmentally-oriented farms (Willock *et al.*, 1999) and some hobby farms (Daniels, 1986; Holloway, 2002). The existence of non-financial/non-pecuniary motives or potential farming lifestyle values may shape farmers' behaviors (Hallam, 1991; Harrington and Reinsel, 1995; Howley, 2015). The ability to change operated farm size may also depend on 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.*, that all farms have the same probability of changing their size category given their initial size) is not likely to hold.

Previous studies that controlled for farm heterogeneity in modeling structural change using the MCM approach have so far only accounted for observed sources of heterogeneity (see Zimmermann and Heckelei (2012) for a recent example). This appears to be restrictive 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 farmer behaviors, which may relate both to observed and unobserved farm and farmer characteristics.

Mixture models have a long history in the literature and have proven to have several advantages over alternative techniques also used to account for unobserved heterogeneity (Greene

and Hensher, 2003; Garver *et al.*, 2008). While the M-MCM has been widely applied in some other strands of economics, to the best of our knowledge, we are the first to evidence heterogeneity in the farm transition process. Recently, Saint-Cyr and Piet (2016) showed that even a restricted mover-stayer model represents farm size dynamics more efficiently than the homogeneous MCM, as has been evidenced in other strands of economics (Blumen *et al.*, 1955; Frydman *et al.*, 1985; Major, 2008; Cipollini *et al.*, 2012). The present article extends this approach in two directions. First, it allows for more than two types of farms and also relaxes the 'pure stayer' assumption. Second, it develops a non-stationary approach to study the impact of agricultural policies on farm size dynamics. This approach leads to separating the farm population into meaningful clusters with different transition patterns (Vermunt, 2010). Since 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.

Since Lee *et al.* (1965), Lee *et al.* (1970) and MacRae (1977) showed that robustly estimating an MCM from aggregate data is possible, structural change has usually been investigated using aggregate data because such data are often easier to obtain than individual-level data. Structural change determinants are thus usually investigated at the macro level. Focusing on agricultural policies, the impacts of public support on the transition probabilities of farms across size categories are generally investigated using transition probability matrices estimated for the overall population of farms (Huettel and Jongeneel, 2011; Zimmermann and Heckelei, 2012; Ben Arfa *et al.*, 2015). Even when individual level data are used, yearly transition probability matrices are first computed for the overall population of farms, the effects of exogenous variables on transition probabilities being estimated in a second step (Hallberg, 1969; Ethridge *et al.*, 1985; Rahelizatovo and Gillespie, 1999). In doing so, individual effects of farms cannot be taken into account. The discrete choice approach adopted here to model farm transitions enables individual farm effects on structural change to be more easily incorporated.

The article is structured as follows. The following section presents the proposed model, the specification method for transition probabilities and the estimation procedure. Then, methods for assessing and interpreting the model's results are presented. The next two sections report the application to a panel of French farms, starting with a description of the data used and explanatory variables investigated, followed by a presentation of the results. Lastly, concluding remarks are made with some considerations given to possible improvements of this study for further research.

2 The non-stationary mixed-Markov chain model (M-MCM)

Let N be the total number of farms in the population and K the total number of farm size categories or choice alternatives. As farm size categories are observed at discrete times, generally 1-year intervals, a discrete-time process is assumed. Denote by y_{it} the size category of a specific

farm i ($i \in N$) at time t ($1 \leq t \leq T$). The indicator variable y_{i1} is equal to $j \in \{0, 1, 2, \cdots, K\}$ if farm i is in category j at time $t = 1$. The category $j = 0$ may indicate entry into or exit from farming. Since farms may enter or leave the farming sector at different dates, the length of time T_i a farm is observed may vary across farms (*i.e.*, $T_i \leq T$). Over the time period T_i, the size evolution of a specific farm i can therefore be represented by the vector $\mathbf{y}_i = (y_{i1}, y_{i2}, \cdots, y_{iT_i})$. Farm movements across size categories are generally supposed to follow a first-order Markov process (Zimmermann *et al.*, 2009). It is thus assumed that the size category of farm i at any time t (y_{it}) depends only on its immediately preceding state (*i.e.*, its size category 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 of each other (Dias and Willekens, 2005). Then, the probability function describing farm movements across size categories can be derived as:[1]

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

where $P(y_{it}|y_{it-1})$ is the transition probability; that is, the probability that farm i chooses a specific size category at time t given its state at time $t-1$.

Suppose now that the observed sample of farms is divided into G homogeneous types instead of just one, each type grouping farms with a similar transition process. The density function of $\mathbf{y}_i$, as a discrete mixing distribution with G support points, can be rewritten as (McLachlan and Peel, 2001):

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

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 to one. In statistics, π_g is called the mixing distribution and $f_g(\mathbf{y}_i)$ is called the mixed function (Lindsay and Lesperance, 1995). Since we defined a finite number of farm types, the mixed model can also mbe called a 'latent class model' with G latent transition processes (Garver *et al.*, 2008; Train, 2009). The density function of $\mathbf{y}_i$ is thus conditional on the mixing distribution and we can represent farm size dynamics as:

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

From equation (3), it can be seen that under the M-MCM, farm size evolution depends on two sets of probabilities. The first term represents the probabilities that farm i belongs to a specific farm type g while the second term represents the probabilities of making transitions across size categories given that farm i belongs to type g.

[1] It should be noted that this formulation does not account for the entries into nor exits from farming. This formulation is used because of data limitations that will be discussed further.

2.1 Specification of the model

Both parts of equation (3) can be specified as a function of exogenous variables (Vermunt, 2010). In this case, the model is fully parametric. However, it has been shown that the resulting estimates could depend on the specification of the mixing distribution (Heckman and Singer, 1984); and the models' parameters may be biased when mis-specifications occur (Wedel, 2002). Furthermore, a parametric approach is convenient when performing discriminant analysis; that is, if we know *a priori* the total number of types in the population that depend on socio-economic characteristics (Garver *et al.*, 2008). Since we have no information about farm types, a semi-parametric approach is applied here, and only transition probabilities are specified as a function of exogenous variables.

As size categories of farms are finite, exhaustive and mutually exclusive, a discrete-choice approach is used to specify transition probabilities. The discrete choice approach assumes that a farmer's choice of the initial size category as well as the choice of making some transitions across categories can be represented by a random utility model (Train, 2009). In our case, farmer's utility may represent the net benefit that arises from choosing (*i.e.*, moving to) a specific size category given the preceding state.

Denote by U_{ijkt} the utility of farm i that results from moving from category j to category k between time periods t and $t+1$. Under the standard behavioral assumption, it is supposed that farmers choose the category which maximizes 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 for two consecutive dates are considered as making a transition from j to j. Under a mixture assumption, the utility level of a farm is conditional on its specific type g. A farm belonging to type g will thus make a transition from a specific category j to category 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 conditional on belonging to type g.

As the utility of farmers may vary over time due to changes in their social and economic environment, transition probabilities are specified as a function of a set of causal factors $\mathbf{x}$. Under the mixture assumption, the utility that would accrue to farm i upon moving from category j to category k at time t given 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 the considered explanatory variables, $\boldsymbol{\beta}_g$ are the parameters to be estimated and $\epsilon_{ijkt|g}$ is an iid random error term specific to farm type g. Explanatory variables are lagged 1-year since farmers' decisions on entering or leaving farming as well as on expansion or contraction, are likely to depend on information available during the previous period (Zepeda, 1995b). According to equation (4), the utility function of farmers thus consists of an observed part ($V_{ijkt|g}$) which depends on some observable individual farm level exogenous variables and a

stochastic component ($\epsilon_{ijkt|g}$) which is the difference between the utility level that farm i experiences and the representation obtained from the chosen explanatory variables (Train, 2009). Since farmers face multiple choices at each occasion, it is econometrically convenient to use a multinomial specification (Greene, 2006). This specification assumes that the error terms $\epsilon_{ijkt|g}$ are randomly drawn from a Gumbel distribution. This leads to the logit multinomial expression for the conditional probability of making a transition from category j to category k at date t:

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

where $\boldsymbol{\beta}_{jk|g}$ is a vector of parameters specific to each farm type g and each jk transition. Choosing staying in the same category for two consecutive years as the reference leads to setting $\boldsymbol{\beta}_{jj|g} = \boldsymbol{0} \quad \forall g = 1, 2, \cdots, G$ and $\forall j = 1, 2, \cdots, K$ for identification purposes. This constraint leads to computing $K - 1$ logarithmic odds ratios for each initial size category j as:

$$\ln\left[\frac{P(y_{it} = k|y_{it-1} = j, g_i = g, \mathbf{x}_{it-1})}{P(y_{it} = j|y_{it-1} = j, g_i = g, \mathbf{x}_{it-1})}\right] = (\boldsymbol{\beta}'_{jk|g} - \boldsymbol{\beta}'_{jj|g})\mathbf{x}_{it-1} \tag{6}$$

2.2 Estimation procedure

The parameters of the model are estimated using the maximum likelihood estimation method. Let $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n)$ be the observed random sample which is composed of a mixture of G types of farms, where the vector $\mathbf{y}_i = (y_{i1}, y_{i2}, \cdots, y_{iT_i})$ collects farm i states over time period $T_i \leq T$. Denote the conditional transition probability of moving from category j to category k by:

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

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

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

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 indicator d_{ijkt} is equal to 1 if farm i moves from category j to category k at time t (*i.e.*, $y_{it} = k|y_{it-1} = j$) and zero otherwise. Since the farm type is unknown beforehand and given some numerical difficulties associated with the maximization of the above expression, the expectation-maximization (EM) algorithm is generally used to estimate the parameters of such models (McLachlan and Krishnan, 2007). The EM algorithm developed by Dempster *et al.* (1977) simplifies the complex log-likelihood in equation (7) into a set of easily solvable

log-likelihood functions by introducing a so-called 'missing variable'.[2]

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 $\mathbf{y}_c = (\mathbf{y}, \mathbf{v})$, is therefore:[3]

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

where $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_G)$ is the vector gathering shares of farm types which has also to be estimated. The complete log-likelihood is thus obtained by:

$$LL_c(\boldsymbol{\beta}, \boldsymbol{\pi}) = \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig} \ln \left\{ \pi_g \prod_{t=1}^{T_i} \prod_{j,k}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right]^{d_{ijkt}} \right\} \tag{9}$$

In this case, v_{ig} is called the 'posterior' probability that farm i belongs to type g given that $\mathbf{y}_i$ has been observed, that is $P(v_{ig} = 1|\mathbf{y}_i)$, while π_g is the 'prior' probability of the mixture (McLachlan and Peel, 2001). This log-likelihood can then be split into two components:

$$\begin{aligned} LL_1 &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig} \ln \pi_g \\ LL_2 &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \ln \left\{ P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \right\} \end{aligned} \tag{10}$$

Since 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 of the four following steps:

(i) Initialization: Arbitrarily choose initial values $\boldsymbol{\Psi}^0 = (\boldsymbol{\Phi}_1^0, \boldsymbol{\Phi}_2^0, \cdots, \boldsymbol{\Phi}_G^0)$ 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 $\boldsymbol{\beta}_{jj|g}^0$ set to zero for identification as previously mentioned.

(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 $\boldsymbol{\Psi}^p$. This conditional expectation probability, that is, the posterior probability $v_{ig}^{(p+1)} = v_{ig}(\mathbf{y}_i; \boldsymbol{\Psi}^p)$, can be obtained

[2] Indeed, the likelihood does not yield an explicit solution for the model parameters. The EM algorithm enables the maximization of a log of sums to be transformed into a recursive maximization of the sum of logs (McLachlan and Krishnan, 2007).

[3] This assumption means that the distribution of the complete-data vector implies the appropriate distribution for the incomplete-data vector (McLachlan and Peel, 2001).

according to Bayes' Law:

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

Replacing v_{gi} in equation (9) by its expected value $v_{ig}^{(p+1)}$ leads to the conditional expectation of the complete data log-likelihood.

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

$$\boldsymbol{\beta}^{(p+1)} = argmax_{\boldsymbol{\beta}} \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \ln \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g})\right] \tag{12}$$

The maximization process of the above equation is straightforward. The transition probability parameters ($\boldsymbol{\beta}^p$) are updated considering $v_{gi}(\mathbf{y}_i; \boldsymbol{\Psi}^p)$ as a weighting factor for each observation (Pacifico and Yoo, 2012). Given equation (5), the log-likelihood function for the transition probabilities can be rewritten as:

$$\begin{aligned} LL_2(\boldsymbol{\beta}) &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \ln \left[\frac{\exp(\boldsymbol{\beta}_{jk|g}^{'} \mathbf{x}_{it-1})}{\sum_{l=1}^{K} \exp(\boldsymbol{\beta}_{jl|g}^{'} \mathbf{x}_{it-1})}\right] \\ &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \left\{ (\boldsymbol{\beta}_{jk|g}^{'} \mathbf{x}_{it-1}) - \ln \left[\sum_{l=1}^{K} \exp(\boldsymbol{\beta}_{jl|g}^{'} \mathbf{x}_{it-1})\right] \right\} \end{aligned} \tag{13}$$

The first-order derivative of this log-likelihood function then becomes:

$$\begin{aligned} \frac{\partial LL_2(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_{jk|g}} &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j}^{K} d_{ijkt} \mathbf{x}_{it-1} - \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j,k}^{K} d_{ijkt} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \mathbf{x}_{it-1}\right] \\ &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j=1}^{K} d_{ijkt} \mathbf{x}_{it-1} - \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j}^{K} \left[P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g}) \mathbf{x}_{it-1}\right] \\ &= \sum_{i=1}^{N} \sum_{g=1}^{G} v_{ig}^{(p+1)} \sum_{t=1}^{T_i} \sum_{j=1}^{K} \left[d_{ijkt} - P(\mathbf{x}_{it-1}; \boldsymbol{\beta}_{jk|g})\right] \mathbf{x}_{it-1} \end{aligned} \tag{14}$$

The model's parameters are obtained by solving the equation $\partial LL_2(\boldsymbol{\beta}) / \partial \boldsymbol{\beta}_{jk|g} = 0 \; \forall j \in K$; $k = 1, 2, \cdots, K-1$; and $g = 1, 2, \cdots, G$.
Then, the probabilities of belonging to farm type g are updated as follows :

$$\pi_g^{(p+1)} = \frac{\sum_{i=1}^{N} v_{ig}^{(p+1)}}{\sum_{i=1}^{N} \sum_{h=1}^{G} v_{ih}^{(p+1)}}, \qquad \forall g \in G \tag{15}$$

(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 (7). At convergence, the resulting parameters are considered to be the optimal estimators ($\hat{\boldsymbol{\Psi}}$).

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 (7) does not allow for the identification of a global maximum in the mixture model, even for discrete mixtures of multinomial logit (Hess *et al.*, 2007). Given the potential presence of a number of local maxima, the EM solutions may be highly dependent on the initial values chosen for $\boldsymbol{\Psi}^0$. Various techniques have been used in the literature to avoid suboptimal solutions (see Baudry and Celeux (2015) for a short review). In the present study, the EM algorithm was run several times with various randomly chosen initial values and those providing the largest likelihood at convergence were chosen as the best ones.

3 Farm types and elasticities of explanatory variables

3.1 Choosing the optimal number of farm types

The total number of components for a mixture model can be chosen either by applying an *a priori* assumption or via information criteria. In the latter case, selection criteria are generally based on the value of $-2LL_G(\mathbf{y};\hat{\boldsymbol{\Psi}})$ of the corresponding model, where matrix $\hat{\boldsymbol{\Psi}}$ 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 preferred (Andrews and Currim, 2003). The selection criteria derive from the following formula:

$$C_G = -2\left\{LL_G(\mathbf{y};\hat{\boldsymbol{\Psi}})\right\} + \kappa N_G \tag{16}$$

where $LL_G(\mathbf{y};\hat{\boldsymbol{\Psi}})$ 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. Different values of κ lead to the two well-known information criteria: the Akaike Information Criterion (AIC) with $\kappa = 2$ and the Bayesian Information Criterion (BIC) with $\kappa = \log N$ with N the total number of observations. Other information criteria can also be derived such as the Consistent Akaike Information Criterion (CAIC) with $\kappa = \log N + 1$ and the modified AIC (AIC3) with $\kappa = 3$.

There is no consensus in the literature about the specific criteria to use for choosing 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;

Sarstedt, 2008).[4]

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 evaluated in terms of elasticities. The 'probability elasticities' measure the effect of a 1% change in the ith explanatory variable on the transition probabilities (Zepeda, 1995a). 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:

$$\boldsymbol{\delta}_{jkt|g} = \frac{\partial p_{jkt|g}}{\partial \mathbf{x}_{t-1}} \times \frac{\mathbf{x}_{t-1}}{p_{jkt|g}}, \qquad \forall j,k \in K \quad \forall g \in G \tag{17}$$

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 conditional on belonging to type g. The first term of equation (17) thus represents the marginal effects of explanatory variables and is given by (Greene, 2006):

$$\frac{\partial p_{jkt|g}}{\partial \mathbf{x}_{t-1}} = p_{jkt|g}\left(\boldsymbol{\beta}_{jk|g} - \sum_{l=1}^{K} \boldsymbol{\beta}_{jl|g} p_{jlt|g}\right) \tag{18}$$

where $\boldsymbol{\beta}_{jk|g}$ is the vector of estimated parameters.
Replacing the marginal effects in equation (17) leads to expressing the transition-probability elasticities as:

$$\boldsymbol{\delta}_{jkt|g} = \left(\boldsymbol{\beta}_{jk|g} - \sum_{l=1}^{K} \boldsymbol{\beta}_{jl|g} p_{jlt|g}\right) \mathbf{x}_{t-1} \tag{19}$$

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

$$\boldsymbol{\delta}_{jjt|g} = \left(-\sum_{l=1}^{K} \boldsymbol{\beta}_{jl|g} p_{jlt|g}\right) \mathbf{x}_{t-1}, \qquad \forall g \in G \tag{20}$$

3.3 Farm structure elasticities

Yearly 'structure elasticities' are also derived to measure the impacts of the exogenous variable on the distribution of farms across size categories. Farm structure elasticities measure the per-

[4]It should be noted that based on Monte Carlo simulations the AIC3 is the best segment-retention criterion for a large variety of multinomial data configurations (Andrews and Currim, 2003) as well as when using Markov modeling approaches (Dias and Willekens, 2005). The choice of a good criterion is based on the stability of the resulting estimated parameters and the standard errors variation.

centage 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, 1995a).
Under the mixture modeling framework, the total number of farms in a specific category k at time t can be obtained as:

$$n_{kt} = \sum_{j=1}^{G} \pi_g \sum_{j}^{K} n_{jt-1} p_{jkt|g}, \qquad \forall k \in K, \quad \forall t \in T \tag{21}$$

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

$$\boldsymbol{\eta}_{kt} = \frac{\partial n_{kt}}{\partial \mathbf{x}_{t-1}} \times \frac{\mathbf{x}_{t-1}}{n_{kt}} \tag{22}$$

In equation (21), only transition probabilities depend on exogenous variables ($\mathbf{x}_{t-1}$). Farm structure elasticities can therefore be obtained using the corresponding probability marginal effects in equation (18). At any specific time t, farm structure elasticities are thus derived as:

$$\boldsymbol{\eta}_{kt} = \left(\sum_{g=1}^{G} \pi_g \sum_{j=1}^{K} n_{jt-1} \frac{\partial p_{jkt|g}}{\partial \mathbf{x}_{t-1}} \right) \frac{\mathbf{x}_{t-1}}{n_{kt}} \tag{23}$$

where the marginal effects at the means of the corresponding explanatory variables are replaced by their values derived in the previous subsection.

4 Empirical application

4.1 Data

For the empirical application, an unbalanced panel from the 'Réseau d'Information Comptable Agricole' (RICA) database from 2000 to 2013 is used. RICA is the French implementation of the Farm Accountancy Data Network (FADN). 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. In particular, this database provides information about the total subsidies received by farms from the 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, years. Some farms remained in the database over the whole 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).
This empirical study concentrates on farm size 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°1242/2008, European farms are classified into fourteen economic size (ES) categories, evaluated in terms of total standard output (SO) expressed in Euros[5]. As mentioned before, in France, RICA focuses on 'medium and large' farms, those with SO greater than or equal to 25,000 Euros; this corresponds to ES category 6 and above. 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°1242/2008, the two last categories correspond to 'large farms'. Here, these farms are separated into two classes so that, in the following, they are referred to as, respectively, 'large' and 'very large' farms. 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. Therefore, a constant population is assumed and only transitions between size categories as defined above are investigated.
For estimation purposes, the sample is restricted to farms which were present in the database for at least two consecutive years, in order to observe at least one transition. The corresponding unbalanced panel then comprises 13,786 farms out of the 16,2171 farms in the original database (85.01%), leading to 92,626 (farm$\times$year) observations and 78,840 individual 1-year transitions (including staying in the same category) from 2000 to 2013. Table 1 shows that farms are more likely to remain in their initial size category for two consecutive years than to move to another category. More than 90% of the farms remain in their initial size category whatever the category considered. Finally, it should be noted that remaining in the initial category does not mean that farms do not increase or decrease their size, but that the change is not sufficient to move to a different category as defined in this study.
To ensure convergence of the EM algorithm, especially when a large number of homogeneous farm types is considered, it is further assumed that farms do not move from medium to very large size and vice versa, because only a few movements between these two extreme size categories are observed (see Table 1). This is a common procedure when using the Markov chain model

[5]The standard output of an agricultural product (crop or livestock) is the average monetary value of the agricultural output at farm-gate price, in euro per hectare or per head of livestock. There is a regional SO coefficient for each product, as an average value over a reference period (5 years, except for the SO 2004 coefficient calculated using the average of 3 years). The sum of all the SO per hectare of crop and per head of livestock in a farm is a measure of its overall economic size, expressed in euro (see `http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Standard_output_(SO)`)

to study structural change in farming (see Ben Arfa *et al.* (2015) for a recent example).

4.2 Explanatory variables

The literature provides a large number of factors that may affect farm structural change. These factors have been divided into several categories, but it is generally supposed that farm structural change may be mainly influenced by population and economic growth, technology improvement, the market structure for inputs as well as for outputs, human capital and some agricultural policies or public programs (see Goddard *et al.* (1993); Harrington and Reinsel (1995); Chavas (2001); Eastwood *et al.* (2010) for detailed discussions). Given the multiplicity of these factors and the potential interactions that exist among some of them, empirical studies generally focus on determining the impact of specific drivers of structural change. This study aims at investigating the impacts of agricultural policies on farm structural change in France. Subsidies received by farms are thus used to explain transition probabilities. Considering all farms together, this study analyzes the impacts of public support programs originating from the CAP of the EU.

Theoretically, no consensus has been found on the effects of such public programs on structural change in farming (Zimmermann and Heckelei, 2012). The impact of such public support may depend on policy objectives and implementation schemes (Goddard *et al.*, 1993). Public farm supports from the CAP are divided into two main components called 'pillars'. While Pillar One aims at improving farmers' competitiveness through, in particular, income support, Pillar Two is dedicated to promoting rural development with broader environmental goals (Gay *et al.*, 2005; Jaime *et al.*, 2016). Despite some reforms aimed at promoting more sustainable agriculture, Pillar One has remained the dominant part of the CAP, accounting for 89% of the CAP budget in 2003 (Gay *et al.*, 2005). Still targeting better performance of public supports, the 2013 reform of the CAP introduced some new regulations such as allowing for the transfer of up to 15% from Pillar One to Pillar Two (European Union, 2013). In the following, we analyze the impacts of farm supports from the CAP focusing on some specific types of subsidies from both pillars.

For Pillar One, direct payments to farms, the most important element of the EU's agricultural policy, are considered. Direct payments are divided into coupled subsidies and decoupled subsidies (single farm payments) and accounted for 72% of CAP expenditure in the 2012 budget (Matthews, 2011). Both kinds of farm subsidies are supposed to encourage farm growth since they may relax liquidity and credit constraints (Roberts and Key, 2008; O'Toole and Hennessy, 2015). Despite the fact that single farm payments were introduced to reduce the distorting impacts of public supports, it has been proven that this kind of farm support may have ex-post coupled effects (Latruffe *et al.*, 2010). Single farm payments may influence farmers' decisions through various channels. In particular, the authors argued that decoupled subsidies influence farmers' decisions through expectations about future payments since such subsidies are often based on historical behavior (Bhaskar and Beghin, 2009). Focusing on direct payments in-

cluding both coupled and decoupled farm subsidies, Pillar One is thus expected to affect farm growth positively.

Considering Pillar Two, we grouped together compensatory allowances for less favored areas (LFA) and for agri-environmental measures (AEMs), and subsidies dedicated to farm investment. While supports for farm investment and agricultural infrastructure can help shift farming practices away from more damaging or wasteful forms of production and towards greater resource efficiency, AEMs are supposed to facilitate change in farming systems toward more resilient production types, better able to cope with future climate-related stress (Dwyer, 2013). These types of support program are more likely to inhibit farm growth. Focusing on these three types of farm support, Pillar Two is expected to have a positive impact on the probability farms remain in the same size category over time.

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

Following Zimmermann and Heckelei (2012), the total initial stock is used as a proxy for path dependency. Initial stock is supposed to negatively affect farm size decline since high initial stock is assumed to result from former investment. Gross Operating Surplus (GOS) minus the total amount of subsidies received and the debt rate of farms are used to reflect the economic environment of farms. The GOS divided by the total number of non-salaried Annual Working Units (AWU) measures the financial capacity of a farm and as such is a very important criteria to obtain credit from a bank for new investments. It could also relate to the self-financing capacity of farms. A positive effect of the GOS is thus expected on the probability of a farm to grow. Debt rate is also expected to have a positive impact on farm growth since credit generally gives firms the resources necessary to expand (Goddard *et al.*, 1993).

Finally, some farm and farmer characteristics such as age, education, legal status, localization, and type of farming, are also used as control variables in the specification of transition probabilities. Such farm and farmer characteristics may allow for observed heterogeneity to be controlled for and are introduced in the model specification using dummy variables. Table 2 presents the description and summary statistics for all the chosen explanatory variables.

Due to potential sources of unobserved heterogeneity as discussed in the general introduction, we further expect that impacts of explanatory variables, especially farm subsidies, may differ from type to type. The results are presented and discussed in the following section.

5 Results

5.1 Type membership and transition probability matrices

According to the information criteria presented above, a mixture of two types of farms seems to be the most appropriate data generating process. Two is chosen as the optimal number specifically for two reasons: first, the BIC and CAIC criteria indicate that two is the optimal number of farm types; second, even if the AIC and AIC3 criteria indicated a higher optimal number of farm types, the results show that the improvement in 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 increases the total number of parameters to be estimated more greatly than the representativeness of the data generating process. For the sake of parsimony, 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 about 68% of the sample consists of farms which tend to remain in their initial size category almost indefinitely, that is, at least during the entire period of observation. We shall therefore name this type the 'almost stayers'. The probabilities of remaining in the same size category for two consecutive years for these farms are close to or over 0.99 (see Table 3a)). This result means that these farms have about a 99% chance of remaining in the same category during a long time period. Conversely, farms belonging to the second type, about 32% of the sample, are more likely to change category in two consecutive years than farms in the first type. We call these 'likely movers'. The transition probability matrix for this farm type is also strongly diagonal meaning that even farms that are likely to change category also have a high probability of remaining in their initial category for two consecutive years (Table 3b)). However, the probability of remaining in the same category for these farms is around 0.85 meaning that they are about 14% more likely to change their size category than those in the 'almost stayer' type.

Summary statistics for various farm and farmer characteristics for both the 'almost stayers' and 'likely movers' farm types were computed in order to identify the profile of farms in each type. Table 4 shows that the probability of belonging to a specific type does not correlate to the farm and farmer characteristics considered in the model specification. There is no significant difference in the distribution of these observed characteristics between the two types of farms. This result means that the unobserved heterogeneity cannot be sufficiently controlled for by the observed farms and/or farmer characteristics considered, and accounting for both kinds of heterogeneity may therefore lead to more efficiently estimating 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 re-

sulting 1-year TPM for the overall population is reported in Table 5. The transition probabilities show that, overall, farms are more likely to remain in their initial category for two consecutive years which is a common feature in agricultural economics (see Hallberg (1969); Stokes (2006) for examples). Indeed, the overall population TPM is strongly diagonal. The probability of remaining in the same size category for two consecutive years is around 0.94. This high probability of remaining in the initial category is due to the high proportion of the 'almost stayer' type in the population. As a consequence, considering a 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 mixed approach may be more informative through the separation of the impacts of explanatory variables, including agricultural policies, on different farm patterns.

5.2 Impact of explanatory variables

We estimated both the mixture model and a homogeneous model (*i.e.*, without accounting for unobserved farm heterogeneity). The results confirm our expectation that the impacts of most explanatory variables depend on the type of farm considered: even if the same sign is most often observed, the coefficient values are generally significantly different. The estimated coefficients for the odds ratios of transition probabilities for both models along with z-score tests for the comparison of the estimated parameters are reported in Tables 6 to 9. Here, an odds ratio is the ratio of the probability that a farm chooses a specific size category over the reference choice of remaining in the same category for two consecutive years. The values of the statistic lead to a rejection of the null hypothesis with a p-value of 5% in most cases. As the values of the odds-ratios coefficients are difficult to interpret, only transition probability elasticities for the main explanatory variables are discussed in the following.

Tables 10, 11 and 12 respectively report the probability elasticities for farms to grow, to decline and to remain in the same size category for two consecutive years for both the homogeneous model and the mixture model. The first two columns of the tables present the results obtained using the homogeneous model. The four following columns present the results for the mixture model considering, respectively, the 'almost stayers' type and the 'likely movers' type. The results confirm our expectations for almost all explanatory variables. The results from both models show that, overall, subsidies from the two pillars of the CAP, initial stock, total GOS and the debt rate of farms have a positive impact on the probability of growing and a negative impact on the probability of shrinking. These results are also consistent with the literature. Indeed, several studies have shown that farm subsidies from support programs of the CAP are likely to favor farm growth (see Ben Arfa *et al.* (2015) for a recent example). Zimmermann and Heckelei (2012) found that initial stock positively affects farm growth. In addition, the present results also support the hypothesis that initial stock decrease the probability of farms to decline. The results from both models also show that, overall, farms are more likely to decrease in size

when farmers are aged over 55 years. Individual legal status as well as specialization in crop productions have the same impacts on farm size. These results can be explained in three ways. First, older farmers may be less motivated to increase their capacity of production when they are close to retirement because they may be more interested in the farm succession (Potter and Lobbley, 1992). In contrast, younger farmers may be more likely to seek to increase agricultural activity, as they are likely to be less financially secure than their older counterparts (Howley *et al.*, 2014). Second, individual legal status farms may face more economic constraints (capital, access to credit, *etc.*) than corporate farms, which may hamper new investments, or such farms may just have less financial motives than corporate farms (Boehlje, 1992). Thirdly, in France, it may be more difficult for farms specialized in crop production to increase their economic size because of the regulation of the land market, since increasing the production capacity for crop farms would generally mean increasing the operated area contrary to livestock breeding farms which can increase the number of animals raised. Conversely, the probability of growing increases if most of the farm is located in an area without natural constraints or if the farmer has a minimum agricultural education. The latter result confirms the positive impact of the managerial capacity on farm size growth (Boehlje, 1992; Goddard *et al.*, 1993).

The resulting elasticities are different considering the homogeneous model and a mixture of 'almost stayers' and 'likely movers' farms. Overall, the impacts of explanatory variables are larger on the transition probabilities of 'likely movers' farms than on those of 'almost stayers' ones, except for farm subsidies. In most cases, the impact of Pillar One is higher for 'almost stayers' farms. The impact of subsidies from Pillar Two on the probability of farms to grow is higher for farms belonging to the 'almost stayer' type than for farms in the 'likely movers' one. For example, a 1% increase in the amount of Pillar One subsidies is expected to increase the probability of farms moving from the medium to the large size by about 0.84% for the 'almost stayers' but only 0.34% for the 'likely movers'. The impact of coupled and decoupled subsidies is thus more than twice higher for 'almost stayers' than for 'likely movers'. This result may be explained by the fact that the 'almost stayers' type may group together farms with liquidity constraints, and support from the CAP may therefore have an income multiplier effect for such farms (Latruffe *et al.*, 2010).

Nevertheless, the results also show that, for the 'almost stayers' type, Pillar One subsidies only have a positive impact on growth for medium sized farms. Indeed, large and very large farms in this type would be more likely to remain in their size category if their amount of subsidies were to increase. It could be that farms in the 'almost stayers' type may be less motivated to increase their operated farm size to a certain level because of, for example, less or non-pecuniary motives (Harrington and Reinsel, 1995). In this case, they may have a lower optimal size than farms in the 'likely movers' type (Howley *et al.*, 2014). As the subsidies received will increase total profit, the probability for these farms to become very large may decrease with the amount of subsidies received.

The results from both the homogeneous model and the mixture of 'almost stayers' and 'likely

movers' also show that Pillar One has higher effects on farm size change than Pillar Two supports. Considering the 'almost stayers' type of farms, the impact of Pillar One on the probability of growing is about six times higher than that the impact of Pillar Two. This result confirms the expectation and previous findings of the literature that subsidies from Pillar One of the CAP are more likely to encourage farm growth than subsidies from Pillar Two (Piet *et al.*, 2012).
Structure elasticities were then computed at the mean point for farm subsidies from both pillars in order to assess their impact on the distribution of farm size. As entries and exits could not be considered due to data limitations, the structure elasticities were computed on farm proportions by size category. The structure elasticities thus represent the variation of the share of farms in a specific category that would occur if the total amount of subsidies received were to increase by 1%. The Figures 2 and 3 presents the resulting structure elasticities for both the homogeneous and the mixture models.
Figures 2 and 3 show that the resulting structure elasticities are quite different from one model to the other. Overall, the homogeneous model tends to overestimate the impacts of subsidies from both pillars. For example, the homogeneous model predicts that a 1% increase of Pillar One subsidies will decrease the proportion of medium sized farms by about 0.06% while the mixture model predicts a decrease by about 0.05% (see Figure 2). Likewise, the homogeneous model predicts a 0.01% increase in the proportion of very large farm if Pillar Two subsidies were to increase by 1% while this proportion is about 0.005% for the mixture model (see Figure 3). These results confirm that ignoring unobserved heterogeneity in modelling farm size dynamics can lead to misleading results. The results also confirm that the impact of subsidies from Pillar One of the CAP on the variation of farm proportion in category of sizes is higher than the impact of supports from Pillar Two. The resulting structure elasticities are about 6 times higher for Pillar One than for Pillar Two whatever the size category considered.

6 Concluding remarks

The aim of this article was to investigate the impacts of a selected set of causal factors likely to influence structural change in farming, accounting for heterogeneity in farmers' behaviors. From a methodological point of view, a non-stationary mixed-Markov chain model (M-MCM) was developed for the first time in agricultural economics. Using individual level data, a mixture of homogeneous farm types was considered in order to incorporate unobserved heterogeneity in the transition process of farms. A discrete choice modeling approach was used to describe the farm size category choices of farmers. A multinomial logit specification for the transition probabilities was implemented and the expectation-maximization (EM) algorithm allowed the estimation of the parameters of the M-MCM. Transition probability elasticities as well as structure elasticities were derived to analyze the impacts of a set of exogenous variables on farm size change in the French farming sector, especially focusing on the impact of specific measures of the CAP of the EU.

Using a sample of farms from the RICA from 2000 to 2013, the results show that French farms can be divided into two unobserved types: 'almost stayers' that are more likely to remain almost indefinitely in their initial size category, and 'likely movers' that change size category more frequently. The results also show that the French farm population consists of a higher proportion of farms that behave like 'almost stayers' leading to a strongly diagonal transition probability matrix for the overall population. Descriptive statistics for both types of farms on a set of observed variables show that the probability of belonging to the 'almost stayer' type or the 'likely movers' type are not significantly correlated with the chosen set of observed farm and/or farmer characteristics, meaning that unobserved heterogeneity cannot be fully controlled for by observed heterogeneity.

The results also show that the impacts of farm subsidies from Pillar One and Pillar Two of the CAP depend on the type that a farm belongs to. Aggregated at the population level, structure elasticities show that an homogeneous and a mixed Markov chain models lead to different results, the homogeneous model tending to either underestimate or overestimate the impacts of subsidies from both pillars on farm size dynamics. These results are relevant for policy assessment since they confirm that ignoring potential unobserved heterogeneity in farmer behavior may lead to inaccurate parameters and therefore to misleading policy recommendations.

However, this study has some limitations that may motivate further research. In the current article, the estimations were performed without taking into account entries into nor exits from the farming sector because of the RICA data limitations at the individual level. Accounting for entries into and exits from farming would constitute an obvious way to analyze the impact of agricultural policy on the total number of farms by size category. The results could be also improved by considering explanatory variables other than those used in this study since it has been proven that several other factors may play an important role on structural change in farming. Lastly, a possible improvement of this work would be to introduce explanatory variables in the specification of the selection of farm types. This could lead to more insightful analyses of the impact of farm and/or farmer characteristics on the probability of belonging to a specific farm type.

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(4):315–321.

Baudry, J.-P. and Celeux, G. (2015). EM for mixtures. *Statistics and Computing*, 25(4):713–726.

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(1):19–42.

Bhaskar, A. and Beghin, J. C. (2009). How coupled are decoupled farm payments? a review of the evidence. *Journal of Agricultural and Resource Economics*, 34(1):130–153.

Blumen, I., Kogan, M., and McCarthy, P. J. (1955). *The industrial mobility of labor as a probability process*. Cornell Studies in Industrial and Labor Relations. 163 p.

Boehlje, M. (1992). Alternative models of structural change in agriculture and related industries. *Agribusiness*, 8(3):219–231.

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.

Bostwick, D. (1962). Yield probabilities as a markov process. *Agricultural Economics Research*, 14(2):49–56.

Bowman, M. S. and Zilberman, D. (2013). Economic factors affecting diversified farming systems. *Ecology and Society*, 18:33(1).

Breustedt, G. and Glauben, T. (2007). Driving forces behind exiting from farming in western europe. *Journal of Agricultural Economics*, 58(1):115–127.

Chavas, J.-P. (2001). Chapter 5 structural change in agricultural production: Economics, technology and policy. In *Handbook of Agricultural Economics*, 1(A): 263–285.

Chavas, J.-P. and Magand, G. (1988). A dynamic analysis of the size distribution of firms: The case of the us dairy industry. *Agribusiness*, 4(4):315–329.

Cipollini, F., Ferretti, C., and Ganugi, P. (2012). Firm size dynamics in an industrial district: The mover-stayer model in action. In Di Ciaccio, A., Coli, M., and Angulo Ibanez, J. M., eds. *Advanced Statistical Methods for the Analysis of Large Data-Sets*, *Springer Berlin Heidelberg*, 443–452.

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(2):121–140.

Daniels, T. L. (1986). Hobby farming in america: Rural development or threat to commercial agriculture? *Journal of Rural Studies*, 2(1):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):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(3):135–157.

Disney, W. T., Duffy, P. A., and Hardy, W. (1988). A Markov chain analysis of pork farm size distributions in the South. *Southern Journal of Agricultural Economics*, 20(2):57–64.

Dwyer, J. (2013). Transformation for sustainable agriculture: what role for the second pillar of cap? *Bio-based and Applied Economics*, 2(1):29–47.

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

Ethridge, D. E., Roy, S. K., and Myers, D. W. (1985). A markov chain analysis of structural changes in the Texas high plains cotton ginning industry. *Southern Journal of Agricultural Economics*, 17(2):11–20.

European Union (2013). *Overview of CAP Reform 2014-2020*. Agricultural Policy Perspectives Brief No.5. 10 p.

Freshwater, D. and Reimer, B. (1995). Socio-economic policies as causal forces for the structure of agriculture. *Canadian journal of agricultural economics/Revue canadienne d'agroeconomie*, (Special Issue):209–222.

Frydman, H., Kallberg, J. G., and Kao, D.-L. (1985). Testing the adequacy of Markov chain and mover-stayer models as representations of credit behavior. *Operations Research*, 33(6):1203–1214.

Gaffney, P. (1997). *A projection of Irish agricultural structure using Markov chain analysis*. CAPRI Working Paper 97-10. 22 p.

Garcia, P., Offutt, S. E., and Sonka, S. T. (1987). Size distribution and growth in a sample of Illinois cash grain farms. *American Journal of Agricultural Economics*, 69(2):471–476.

Garver, M. S., Williams, Z., and Taylor, G. S. (2008). Employing latent class regression analysis to examine logistics theory: an application of truck driver retention. *Journal of Business Logistics*, 29(2):233–257.

Gay, S., Osterburg, B., Baldock, D., and Zdanowicz, A. (2005). *Recent evolution of the EU Common Agricultural Policy (CAP): State of Play and Environmental Potential.* MEACAP WP6 D4b, specific target research project SSPE-CT-2004-503604. 59 p.

Gillespie, J. M. and Fulton, J. R. (2001). A markov chain analysis of the size of hog production firms in the united states. *Agribusiness*, 17(4):557–570.

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(4):475–489.

Greene, W. H. (2006). *Econometric analysis*. Pearson Education India. 1216 p., 6th edition.

Greene, W. H. and Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8):681 – 698.

Hallam, A. (1991). Economies of size and scale in agriculture: An interpretive review of empirical measurement. *Review of Agricultural Economics*, 13(1):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(2):289–302.

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

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2):271–320.

Hess, S., Bierlaire, M., and Polak, J. (2007). A systematic comparison of continuous and discrete mixture models. *European Transport*, 37:35–61.

Holloway, L. (2002). Smallholding, hobby-farming, and commercial farming: Ethical identities and the production of farming spaces. *Environment and Planning A*, 34(11):2055–2070.

Howley, P. (2015). The Happy Farmer: The Effect of Nonpecuniary Benefits on Behavior. *American Journal of Agricultural Economics*, 97(4):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(2):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(4):497–527.

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(1):66–91.

Jaime, M. M., Coria, J., and Liu, X. (2016). Interactions between CAP agricultural and agri-environmental subsidies and their effects on the uptake of organic farming. *American Journal of Agricultural Economics*, 98(4):1114–1145.

Judge, G. G. and Swanson, E. R. (1962). Markov chains: basic concepts and suggested uses in agricultural economics. *Australian Journal of Agricultural Economics*, 6(2):49–61.

Karantininis, K. (2002). Information-based estimators for the non-stationary transition probability matrix: An application to the Danish pork industry. *Journal of Econometrics*, 107(1-2):275–290.

Keane, M. (1991). Changes in the size structure of Irish dairy farms. *Irish Journal of Agricultural Economics and Rural Sociology*, 14:67–74.

Krenz, R. D. (1964). Projection of farm numbers for North Dakota with Markov chains. *Agricultural Economics Research*, 16(3):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(2):351–365.

Lee, T., Judge, G., and Zellner, A. (1970). *Estimating the parameters of the Markov probability model from aggregate time series data*. Amsterdam: North Holland. 254 p.

Lee, T. C., Judge, G. G., and Takayama, T. (1965). On estimating the transition probabilities of a Markov process. *Journal of Farm Economics*, 47(3):742–762.

Lindsay, B. G. and Lesperance, M. L. (1995). A review of semiparametric mixture models. *Journal of statistical planning and inference*, 47(1-2):29–39.

MacRae, E. C. (1977). Estimation of time-varying Markov processes with aggregate data. *Econometrica*, 45(1):183–198.

Major, K. (2008). Income disparities among Hungarian micro-regions: The mover-stayer model. *Acta Oeconomica*, 58(2):127–156.

Matthews, A. (2011). *Post-2013 EU Common Agricultural Policy, Trade and Development. A Review of Legislative Proposals.* International Centre for Programme on Agricultural Trade and Sustainable Development (ICTSD), Issue Paper No 39. 46 p.

Maybery, D., Crase, L., and Gullifer, C. (2005). Categorising farming values as economic, conservation and lifestyle. *Journal of Economic Psychology*, 26(1):59 – 72.

McInerney, N. and Garvey, E. (2004). *Farm structure and agricultural labour*. CAPRI Working Papers. Dept of Economics, NUI Galway. 18 p.

McLachlan, G. and Krishnan, T. (2007). *The EM algorithm and extensions*, volume 382. John Wiley & Sons. 359 p.

McLachlan, G. and Peel, D. (2001). *Finite mixture models*. John Wiley & Sons. 419 p.

Mzoughi, N. (2011). Farmers adoption of integrated crop protection and organic farming: Do moral and social concerns matter? *Ecological Economics*, 70(8):1536 – 1545.

O'Toole, C. and Hennessy, T. (2015). Do decoupled payments affect investment financing constraints? evidence from Irish agriculture. *Food Policy*, 56:67–75.

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 n°6. 15 p.

Padberg, D. I. (1962). The use of Markov processes in measuring changes in market structure. *Journal of Farm Economics*, 44(1):189–199.

Paternoster, R., Brame, R., Mazerolle, P., and Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, 36(4):859–866.

Piet, L., Latruffe, L., Le Mouel, C., and Desjeux, Y. (2012). How do agricultural policies influence farm size inequality? the example of France. *European Review of Agricultural Economics*, 39(1):5–28.

Potter, C. and Lobbley, M. (1992). Ageing and succession on familly farms: The impact on decision-making and land use. *Sociologia Ruralis*, 32(2-3):317–334.

Rahelizatovo, N. C. and Gillespie, J. M. (1999). Dairy farm size, entry, and declining production region. *Journal of Agricultural and Applied Economics*, 31(2):333–347.

Roberts, M. J. and Key, N. (2008). Agricultural payments and land concentration: a semiparametric spatial regression analysis. *American Journal of Agricultural Economics*, 90(3):627–643.

Saint-Cyr, L. D. F. and Piet, L. (2016). Movers and stayers in the farming sector: accounting for unobserved heterogeneity in structural change. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, : Published online, doi: 10.1111/rssc.12196.

Sarstedt, M. (2008). Market segmentation with mixture regression models: Understanding measures that guide model selection. *Journal of Targeting, Measurement and Analysis for Marketing*, 16(3):228–246.

Stokes, J. R. (2006). Entry, exit, and structural change in pennsylvania's dairy sector. *Agricultural Resource and Economics Review*, 35(2):357–373.

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*, 43(3):359 –382.

Vermunt, J. (2010). Longitudinal research using mixture models. In van Montfort, K., Oud, J. H., and Satorra, A., eds. *Longitudinal Research with Latent Variables*, *Springer Berlin Heidelberg*, chap. 4: 119–152.

Wedel, M. (2002). Concomitant variables in finite mixture models. *Statistica Neerlandica*, 56(3):362–375.

Weiss, C. R. (1999). Farm Growth and Survival: Econometric Evidence for Individual Farms in Upper Austria. *American Journal of Agricultural Economics*, 81(1):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(2):286–303.

Zepeda, L. (1995a). Asymmetry and nonstationarity in the farm size distribution of Wisconsin milk producers: An aggregate analysis. *American Journal of Agricultural Economics*, 77(4):837–852.

Zepeda, L. (1995b). Technical change and the structure of production: a non-stationary markov analysis. *European Review of Agricultural Economics*, 22(1):41–60.

Zimmermann, A. and Heckelei, T. (2012). Structural change of European dairy farms: A cross-regional analysis. *Journal of Agricultural Economics*, 63(3):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(5):601–618.

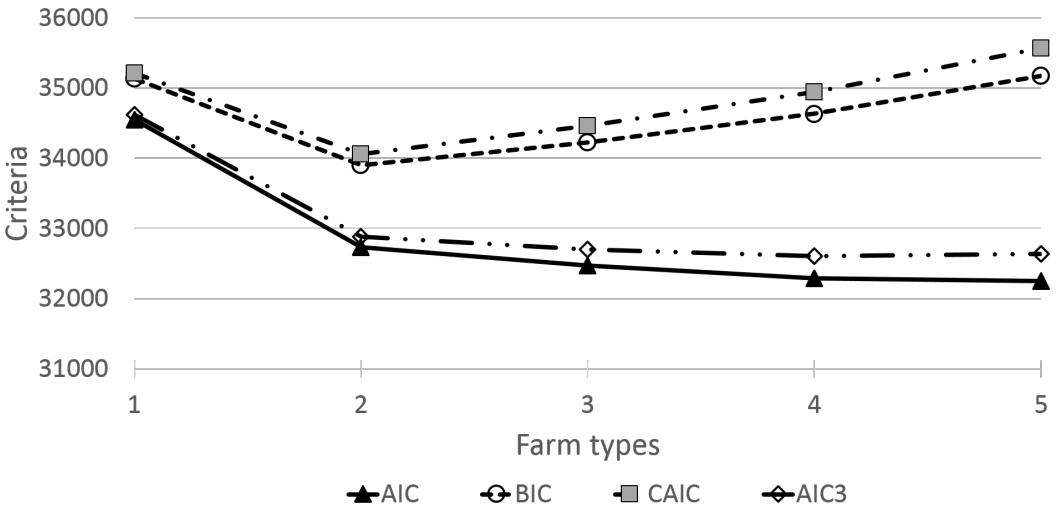
Table 1: Observed 1-Year Transitions across Size Categories, 2000-2013.

		ES class			
		25-100	100-250	$\geq$ 250	Total transitions
ES class	25-100	27,121	1,335	31	28,487
	100-250	1,280	32,822	1,089	35,191
	$\geq$ 250	39	872	14,251	15,162

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

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

Figure 1: Comparison of model-fit statistics for different numbers of farm type.



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

Table 2: Descriptive Statistics and Definition of Explanatory Variables (n=92,626).

Variable	Unit	Mean	Std. Dev.	Min.	Max.
pillar1	1,000 Euros	29.66	29.00	0.00	315.40
Subsidies from Pillar One					
pillar2	1,000 Euros	12.05	21.85	0.00	443.98
Subsidies from Pillar Two					
in_stock	1,000 Euros	99.07	166.28	0.00	3,624.47
Total initial stock					
GOS/ns-awu	1,000 Euros	27.43	45.90	-1,052.02	1,279.47
GOS per non-salaried AWU					
debt_rate	ratio	39.89	28.62	0.00	578.80
Total debt/liabilities					
age_55	dummy	0.19	0.39	0.00	1.00
1 if farmer is over 55 years old[a]					
crops	dummy	0.45	0.50	0.00	1.00
1 if farm is specialised in crops					
individual	dummy	0.51	0.50	0.00	1.00
1 if farm has individual legal status					
LFA	dummy	0.40	0.49	0.00	1.00
1 if farm is located in LFA					
agri_skills	dummy	0.93	0.25	0.00	1.00
1 if farmer has agricultural education					

Less Favored Area (LFA); Gross Output Surplus (GOS)
[a] In order to take into account farmers' anticipation, 55 years old is used instead of 60 which is the minimum age for retirement in France.

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

Table 3: Estimated Farm Type Shares and 1-Year Transition Probability Matrices (TPMs).

		Shares	ES class		
		68.75%	25-100	100-250	$\geq$ 250
ES class	25-100	40.68%	0.990(0.019)	0.010(0.019)	-
	100-250	43.39%	0.07(0.017)	0.986(0.018)	0.007(0.007)
	$\geq$ 250	15.93%	-	0.011(0.019)	0.989(0.019)

a) 'Almost stayers' TPM

		Shares	ES class		
		31.25%	25-100	100-250	$\geq$ 250
ES class	25-100	39.03%	0.872(0.090)	0.128(0.090)	-
	100-250	44.78%	0.083(0.076)	0.822(0.071)	0.081(0.070)
	$\geq$ 250	16.19%	-	0.159(0.097)	0.841(0.097)

b) 'Likely movers' TPM

Note: Economic size (ES) in 1,000 Euros of standard output (SO); standard deviations in parenthesis; transition between extreme size categories (25-100 and $\geq$ 250) are constrained to be zero.

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

Table 4: Descriptive Statistics for Observed Farm Characteristics by Type Membership.

Variable	'Almost stayers'		'Likely movers'	
age_55	0.190	(0.392)	0.189	(0.391)
crops	0.438	(0.496)	0.433	(0.495)
individual	0.534	(0.499)	0.517	(0.500)
LFA	0.407	(0.491)	0.408	(0.491)
agri_skills	0.932	(0.252)	0.937	(0.244)

Note: Less Favored Area (LFA); standard deviations in parentheses.
[a] t-tests for the null hypothesis that the means of two different types are equal.

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

Table 5: Overall Population 1-Year Transition Probability Matrix.

		ES class		
		25-100	100-250	$\geq$ 250
ES class	25-100	0.954	0.046	-
	100-250	0.035	0.934	0.031
	$\geq$ 250	-	0.058	0.942

Note: Economic size (ES) in 1,000 Euros of standard output (SO); standard deviations in parentheses; transition between extreme size categories (25-100 and $\geq$ 250) are constrained to be zero.

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

Table 6: Estimated Parameters of Farm Probabilities to Grow between Two Consecutive Years both for the Homogeneous and the Mixed- Markov Chain Models (MCMs).

Odds ratios	Variables	Homogeneous		'Almost stayers'		'Likely movers'	
p12/p11	*intercept*	-3.872***	(0.184)	-6.352***	(0.330)	-2.602***	(0.172)
	pillar1	0.030***	(0.003)	0.047***	(0.004)	0.022***	(0.002)
	pillar2	0.008***	(0.003)	0.015***	(0.004)	0.006**	(0.002)
	in_stock	0.004***	(0.001)	0.005***	(0.001)	0.006***	(0.001)
	GOS/ns-awu	0.017***	(0.002)	0.017***	(0.003)	0.015***	(0.002)
	debt_rate	0.013***	(0.001)	0.017***	(0.002)	0.013***	(0.001)
	age_55	-0.200**	(0.094)	-0.738***	(0.180)	-0.161*	(0.094)
	crops	-0.454***	(0.085)	-0.410***	(0.135)	-0.452***	(0.077)
	individual	-0.615***	(0.076)	-0.734***	(0.114)	-0.614***	(0.070)
	LFA	-0.504***	(0.080)	-0.708***	(0.129)	-0.491***	(0.072)
	agri_skills	0.249*	(0.145)	0.503**	(0.252)	0.252*	(0.136)
p23/22	*intercept*	-4.916***	(0.221)	-5.668***	(0.345)	-4.050***	(0.226)
	pillar1	0.008***	(0.002)	0.023***	(0.002)	0.015***	(0.002)
	pillar2	0.008***	(0.002)	0.012***	(0.003)	0.002	(0.002)
	in_stock	0.003***	(0.000)	0.001	(0.001)	0.004***	(0.000)
	GOS/ns-awu	0.010***	(0.001)	0.005***	(0.001)	0.013***	(0.001)
	debt_rate	0.015***	(0.001)	0.009***	(0.002)	0.016***	(0.001)
	age_55	-0.089	(0.097)	0.000	(0.155)	-0.134	(0.100)
	crops	-0.584***	(0.089)	0.804***	(0.125)	-1.025***	(0.089)
	individual	-0.387***	(0.082)	-0.207	(0.126)	-0.515***	(0.082)
	LFA	-0.253***	(0.088)	-0.479***	(0.160)	-0.164*	(0.084)
	agri_skills	0.298	(0.186)	0.346	(0.284)	0.310*	(0.183)

Note: Less Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); standard errors in parentheses; ***, ** and * indicate significance at 0.1%, 1% and 5% levels, respectively.

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

Table 7: Z-scores for Testing the Equality of the Farm Types' Odds Ratios Coefficients for Farm Growth.

Odds ratios	Variables	'AS'-'HM'		'LM'-'HM'		'AS'-'LM'	
		score	p_value	score	p_value	score	p_value
p12/p11	*intercept*	3.871	0.000	-2.305	0.021	-5.699	0.000
	pillar1	1.597	0.110	-0.656	0.512	- 2.144	0.032
	pillar2	0.866	0.386	1.262	0.207	0.687	0.492
	in_stock	0.240	0.810	-0.602	0.547	- 0.725	0.468
	GOS/ns-awu	2.521	0.012	0.501	0.616	-2.096	0.036
	debt_rate	-2.648	0.008	0.288	0.773	2.839	0.005
	age_55	0.272	0.786	0.011	0.991	- 0.272	0.786
	crops	-0.873	0.383	0.008	0.994	0.901	0.368
	individual	-1.342	0.180	0.124	0.901	1.466	0.143
	LFA	0.872	0.383	0.015	0.988	-0.875	0.382
	agri_skills	-6.562	0.000	5.038	0.000	-10.079	0.000
p23/p22	*intercept*	-10.772	0.000	2.959	0.003	13.131	0.000
	pillar1	1.343	0.179	-2.159	0.031	-3.026	0.002
	pillar2	-3.302	0.001	1.296	0.195	4.272	0.000
	in_stock	-2.930	0.003	1.933	0.053	4.466	0.000
	GOS/ns-awu	-2.474	0.013	0.744	0.457	3.039	0.002
	debt_rate	0.487	0.626	-0.319	0.750	-0.724	0.469
	age_55	9.035	0.000	-3.504	0.000	-11.898	0.000
	crops	1.195	0.232	-1.094	0.274	-2.040	0.041
	individual	-1.235	0.217	0.739	0.46	1.745	0.081
	LFA	0.144	0.886	0.048	0.962	- 0.108	0.914
	agri_skills	-1.839	0.066	2.737	0.006	3.925	0.000

Note: 'Almost stayers' type (AST), Homogeneous model (HM) and 'Likely movers' type (LMT); Least Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); z-scores for testing the null hypothesis that coefficients of two different types are equal (Paternoster *et al.*, 1998).

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

Table 8: Estimated Parameters of Farm Probabilities to Decline between two consecutive years both for the Homogeneous and the Mixed- Markov Chain Models (MCMs).

Odds ratios	Variables	Homogeneous		'Almost stayers'		'Likely movers'	
p21/p22	*intercept*	-1.957***	(0.195)	-3.593***	(0.373)	-0.633***	(0.184)
	pillar1	-0.019***	(0.001)	-0.016***	(0.003)	-0.017***	(0.001)
	pillar2	-0.014***	(0.003)	-0.068***	(0.008)	-0.012***	(0.002)
	in_stock	-0.008***	(0.001)	-0.029***	(0.002)	-0.006***	(0.001)
	GOS/ns-awu	-0.018***	(0.001)	-0.030***	(0.003)	-0.018***	(0.001)
	debt_rate	-0.005***	(0.001)	0.005**	(0.002)	-0.008***	(0.001)
	age_55	0.292***	(0.081)	1.348***	(0.136)	0.119	(0.086)
	crops	-0.070	(0.076)	-1.222***	(0.146)	0.091	(0.071)
	individual	0.640***	(0.074)	0.562***	(0.128)	0.598***	(0.068)
	LFA	0.567***	(0.074)	0.245*	(0.136)	0.612***	(0.070)
	agri_skills	-0.052	(0.150)	0.942***	(0.296)	-0.358**	(0.140)
p32/p33	*intercept*	-1.552***	(0.240)	-3.296***	(0.351)	0.024	(0.230)
	pillar1	-0.002	(0.001)	-0.000	(0.002)	-0.008***	(0.001)
	pillar2	-0.006**	(0.002)	-0.007*	(0.004)	-0.003	(0.003)
	in_stock	-0.004***	(0.001)	-0.002***	(0.001)	-0.005***	(0.000)
	GOS/ns-awu	-0.007***	(0.001)	-0.008***	(0.002)	-0.008***	(0.001)
	debt_rate	-0.007***	(0.002)	-0.014***	(0.003)	-0.008***	(0.002)
	age_55	0.081	(0.102)	0.823***	(0.163)	-0.008	(0.108)
	crops	-0.231**	(0.093)	-0.102	(0.153)	0.048	(0.083)
	individual	0.342***	(0.098)	1.318***	(0.152)	0.125	(0.093)
	LFA	0.323***	(0.104)	-0.709***	(0.221)	0.257***	(0.097)
	agri_skills	0.154	(0.195)	-0.319	(0.266)	-0.007	(0.176)

Note: Less Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); standard errors in parentheses; ***, ** and * indicate significance at 0.1%, 1% and 5% levels, respectively.

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

Table 9: Z-scores for Testing the Equality of the Farm Types' Odds Ratios Coefficients for Farm Decline.

Odds ratios	Variables	'AS'-'HM'		'LM'-'HM'		'AS'-'LM'	
		score	p_value	score	p_value	score	p_value
p21/p22	*intercept*	0.952	0.341	0.978	0.328	-0.387	0.699
	pillar1	-6.326	0.000	0.600	0.549	6.646	0.000
	pillar2	-8.541	0.000	1.419	0.156	9.378	0.000
	in_stock	-3.768	0.000	0.015	0.988	3.715	0.000
	GOS/ns-awu	3.657	0.000	-1.337	0.181	-4.640	0.000
	debt_rate	6.655	0.000	-1.460	0.144	-7.630	0.000
	age_55	-7.008	0.000	1.559	0.119	8.104	0.000
	crops	-0.529	0.597	-0.418	0.676	0.251	0.802
	individual	-2.091	0.037	0.441	0.659	2.408	0.016
	LFA	3.000	0.003	-1.491	0.136	-3.975	0.000
	agri_skills	-3.885	0.000	4.940	0.000	7.116	0.000
p32/p33	*intercept*	0.577	0.564	-3.298	0.001	-2.987	0.003
	pillar1	-0.243	0.808	0.755	0.450	0.788	0.431
	pillar2	2.453	0.014	-0.802	0.423	-3.221	0.001
	in_stock	-0.642	0.521	-0.492	0.623	0.249	0.803
	GOS/ns-awu	-2.158	0.031	-0.603	0.547	1.766	0.077
	debt_rate	3.860	0.000	-0.597	0.551	-4.251	0.000
	age_55	0.719	0.472	2.242	0.025	0.863	0.388
	crops	5.380	0.000	-1.611	0.107	-6.693	0.000
	individual	-4.230	0.000	-0.464	0.643	4.008	0.000
	LFA	-1.438	0.150	-0.615	0.539	0.979	0.328
	agri_skills	-4.102	0.000	4.734	0.000	7.906	0.000

Note: 'Almost stayers' type (AST), Homogeneous model (HM) and 'Likely movers' type (LMT); Least Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); z-scores for testing the null hypothesis that coefficients of two different types are equal (Paternoster *et al.*, 1998).

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

Table 10: Yearly Probability Elasticities of Farm growth both for the Homogeneous and the Mixed- Markov Chain Models (MCMs).

Probability	Variables	Homogeneous		'Almost stayers'		'Likely movers'	
Medium/Large p(12)	*pillar1*	0.514***	(0.042)	0.841***	(0.066)	0.344***	(0.038)
	pillar2	0.068**	(0.022)	0.135***	(0.033)	0.044*	(0.019)
	in_stock	0.143***	(0.028)	0.178***	(0.021)	0.188***	(0.034)
	GOS/ns-awu	0.172***	(0.017)	0.183***	(0.026)	0.147***	(0.019)
	debt_rate	0.356***	(0.031)	0.512***	(0.046)	0.340***	(0.027)
	age_55	-0.033*	(0.016)	-0.126***	(0.031)	-0.025	(0.015)
	crops	-0.169***	(0.032)	-0.156**	(0.051)	-0.160***	(0.028)
	individual	-0.470***	(0.058)	-0.586***	(0.091)	-0.421***	(0.049)
	LFA	-0.278***	(0.044)	-0.406***	(0.074)	-0.244***	(0.037)
	agri_skills	0.217	(0.127)	0.456*	(0.228)	0.199	(0.107)
Large/V. large p(23)	*pillar1*	0.290***	(0.052)	0.792***	(0.084)	0.475***	(0.044)
	pillar2	0.089***	(0.021)	0.139***	(0.032)	0.032	(0.021)
	in_stock	0.259***	(0.032)	0.069	(0.047)	0.313***	(0.032)
	GOS/ns-awu	0.268***	(0.028)	0.123**	(0.040)	0.343***	(0.028)
	debt_rate	0.611***	(0.054)	0.371***	(0.082)	0.657***	(0.053)
	age_55	-0.017	(0.016)	-0.003	(0.027)	-0.024	(0.017)
	crops	-0.265***	(0.041)	0.377***	(0.058)	-0.450***	(0.040)
	individual	-0.172***	(0.034)	-0.086	(0.051)	-0.249***	(0.035)
	LFA	-0.095**	(0.030)	-0.162**	(0.054)	-0.084**	(0.029)
	agri_skills	0.275	(0.171)	0.319	(0.266)	0.304	(0.161)

Note: Less Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); standard errors in parentheses; ***, ** and * indicate significance at 0.1%, 1% and 5% levels, respectively.

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

Table 11: Yearly Probability Elasticities of Farm Decline both for the Homogeneous and the Mixed- Markov Chain Models (MCMs).

Odds ratios	Variable code	Homogeneous MCM		Mixed-MCM 'Almost stayers'		Mixed-MCM 'Likely movers'	
Large/Medium p(21)	*pillar1*	-0.648***	(0.051)	-0.546***	(0.110)	-0.570***	(0.044)
	pillar2	-0.161***	(0.030)	-0.770***	(0.092)	-0.134***	(0.027)
	in_stock	-0.681***	(0.078)	-2.618***	(0.211)	-0.523***	(0.063)
	GOS/ns-awu	-0.498***	(0.035)	-0.804***	(0.072)	-0.496***	(0.037)
	debt_rate	-0.236***	(0.057)	0.209*	(0.102)	-0.368***	(0.053)
	age_55	0.049***	(0.013)	0.23***	(0.023)	0.020	(0.013)
	crops	-0.025	(0.034)	-0.577***	(0.069)	0.066*	(0.030)
	individual	0.253***	(0.029)	0.226***	(0.051)	0.231***	(0.024)
	LFA	0.188***	(0.024)	0.083	(0.046)	0.192***	(0.021)
	agri_skills	-0.056	(0.137)	0.882**	(0.277)	-0.329**	(0.122)
V. Large/Large p(32)	*pillar1*	-0.069	(0.049)	-0.014	(0.081)	-0.307***	(0.053)
	pillar2	-0.087*	(0.036)	-0.109	(0.064)	-0.043	(0.033)
	in_stock	-0.737***	(0.093)	-0.385**	(0.127)	-0.752***	(0.077)
	GOS/ns-awu	-0.335***	(0.052)	-0.421***	(0.080)	-0.320***	(0.045)
	debt_rate	-0.32***	(0.080)	-0.691***	(0.143)	-0.354***	(0.070)
	age_55	0.014	(0.017)	0.148***	(0.029)	-0.001	(0.016)
	crops	-0.100*	(0.040)	-0.048	(0.072)	0.017	(0.030)
	individual	0.073***	(0.021)	0.303***	(0.034)	0.023	(0.017)
	LFA	0.065**	(0.020)	-0.142**	(0.044)	0.052**	(0.019)
	agri_skills	0.137	(0.172)	-0.297	(0.247)	-0.006	(0.142)

Note: Less Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); standard errors in parentheses; ***, ** and * indicate significance at 0.1%, 1% and 5% levels, respectively.

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

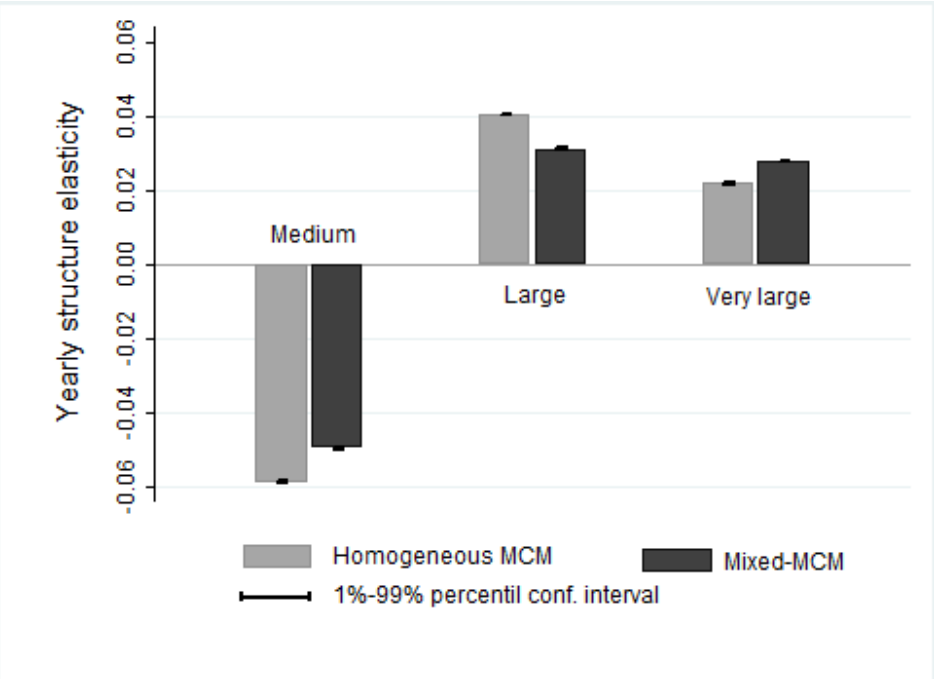
Table 12: Yearly Probability Elasticities of Farms to Remain in the same Category of Sizes both for the Homogeneous and the Mixed- Markov Chain Models (MCMs).

Probability	Variables	Homogeneous		'Almost stayers'		'Likely movers'	
Medium/Medium	*pillar1*	-0.034***	(0.004)	-0.013***	(0.002)	-0.069***	(0.009)
p(11)	*pillar2*	-0.004**	(0.001)	-0.002**	(0.001)	-0.008*	(0.004)
	in_stock	-0.010***	(0.002)	-0.003***	(0.000)	-0.040***	(0.009)
	GOS/ns-awu	-0.013***	(0.002)	-0.003***	(0.001)	-0.034***	(0.006)
	debt_rate	-0.023***	(0.003)	-0.007***	(0.001)	-0.070***	(0.007)
	age_55	0.001*	(0.000)	0.001***	(0.000)	0.003	(0.002)
	crops	0.008***	(0.001)	0.002***	(0.000)	0.024***	(0.003)
	individual	0.018***	(0.002)	0.004***	(0.000)	0.052***	(0.005)
	LFA	0.012***	(0.002)	0.003***	(0.000)	0.033***	(0.004)
	agri_skills	-0.011	(0.006)	-0.005	(0.002)	-0.032	(0.018)
Large/Large	*pillar1*	0.007**	(0.002)	0.005***	(0.000)	-0.009	(0.007)
p(22)	*pillar2*	0.001	(0.001)	0.000	(0.000)	0.009**	(0.003)
	in_stock	0.006**	(0.002)	0.006***	(0.001)	0.003	(0.006)
	GOS/ns-awu	0.000	(0.002)	0.000	(0.001)	-0.005	(0.004)
	debt_rate	-0.015***	(0.003)	-0.004***	(0.001)	-0.036***	(0.008)
	age_55	-0.002*	(0.001)	-0.003***	(0.001)	-0.001	(0.002)
	crops	0.008***	(0.001)	-0.002	(0.001)	0.024***	(0.003)
	individual	-0.011***	(0.002)	-0.002*	(0.001)	-0.027***	(0.005)
	LFA	-0.007***	(0.002)	0.000	(0.000)	-0.026***	(0.005)
	agri_skills	-0.007	(0.008)	-0.008**	(0.003)	0.010	(0.018)
V.Large/V.Large	*pillar1*	0.004	(0.003)	0.000	(0.001)	0.049***	(0.007)
p(33)	*pillar2*	0.005**	(0.002)	0.001*	(0.000)	0.007	(0.005)
	in_stock	0.030***	(0.003)	0.003***	(0.001)	0.085***	(0.007)
	GOS/ns-awu	0.015***	(0.002)	0.003***	(0.000)	0.042***	(0.005)
	debt_rate	0.021***	(0.005)	0.007***	(0.001)	0.065***	(0.012)
	age_55	-0.001	(0.001)	-0.003**	(0.001)	0.000	(0.003)
	crops	0.005**	(0.002)	0.001	(0.001)	-0.003	(0.005)
	individual	-0.007**	(0.002)	-0.008***	(0.002)	-0.006	(0.004)
	LFA	-0.006**	(0.002)	0.001***	(0.000)	-0.012*	(0.005)
	agri_skills	-0.009	(0.012)	0.003	(0.003)	0.001	(0.027)

Note: Less Favored Area (LFA); Gross Output Surplus per non-salaried AWU (GOS/ns-awu); standard errors in parentheses; ***, ** and * indicate significance at 0.1%, 1% and 5% levels, respectively.

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

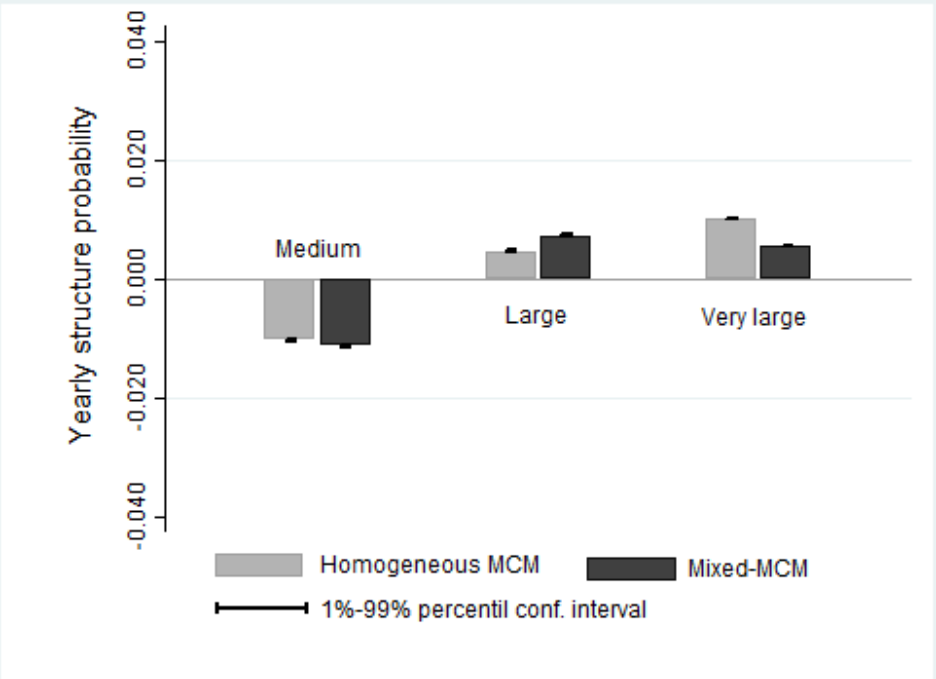
Figure 2: Yearly Structure Elasticities of Farm Subsidies from Pillar One of the CAP both for Homogeneous and Mixed- Markov Chain Models.



Note: 'Medium' sized farms are less than 100,000 Euros of SO; 'Large' farms are 100,000-250,000 Euros of SO; 'Very large' farms are 250,000 Euros of SO and more.

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

Figure 3: Yearly Structure Elasticities of Farm Subsidies from Pillar Two of the CAP both for Homogeneous and Mixed Markov Chain Models.



Note: 'Medium' sized farms are less than 100,000 Euros of SO; 'Large' farms are 100,000-250,000 Euros of SO; 'Very large' farms are 250,000 Euros of SO and more.

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

Les **Working Papers SMART – LERECO** sont produits par l'UMR SMART-LERECO

- **UMR SMART-LERECO**

L'Unité Mixte de Recherche (UMR 1302) *Laboratoire d'Etudes et de Recherches en Economie sur les Structures et Marchés Agricoles, Ressources et Territoires* comprend les unités de recherche en Economie de l'INRA de Rennes et de l'INRA de Nantes et les membres des Unités Pédagogiques de Rennes et Angers du département Economie, Gestion et Société d'Agrocampus Ouest.

Adresse:
UMR SMART-LERECO, 4 allée Adolphe Bobierre, CS 61103, 35011 Rennes cedex

Site internet : https://www6.rennes.inra.fr/smart/

Liste complète des Working Papers SMART – LERECO :
https://www6.rennes.inra.fr/smart/Working-Papers-SMART-LERECO
https://ideas.repec.org/s/rae/wpaper.html
http://ageconsearch.umn.edu/handle/204962/

The **Working Papers SMART – LERECO** are produced by UMR SMART-LERECO

- **UMR SMART-LERECO**

The « Mixed Unit of Research » (UMR1302) *Laboratory for Empirical Research in Economics on Structures and Markets in Agriculture, Resources and Territories* is composed of the research units in Economics of INRA Rennes and INRA Nantes and of the members of the Agrocampus Ouest's Department of Economics, Management and Society who are located in Rennes and Angers.

Address:
UMR SMART-LERECO, 4 allée Adolphe Bobierre, CS 61103, 35011 Rennes cedex

Website: https://www6.rennes.inra.fr/smart_eng/

Full list of the Working Papers SMART – LERECO:
https://www6.rennes.inra.fr/smart_eng/Working-Papers-SMART-LERECO
https://ideas.repec.org/s/rae/wpaper.html
http://ageconsearch.umn.edu/handle/204962/

Contact

Working Papers SMART – LERECO
INRA, UMR SMART-LERECO
4 allée Adolphe Bobierre, CS 61103
35011 Rennes cedex, France
Email : smart-lereco-wp@inra.fr

2017

Working Papers SMART – LERECO

UMR INRA-Agrocampus Ouest **SMART-LERECO** (Laboratoire d'Etudes et de Recherches en Economie sur les Structures et Marchés Agricoles, Ressources et Territoires)

Rennes, France