Short Term Prediction of Agricultural Structural Change using Farm Accountancy Data Network and Farm Structure Survey Data

Kurzfristige Vorherhersage des Agrarstrukturwandels, basierend auf FSS- und FADN-Daten

Hugo Storm and Thomas Heckelei
University of Bonn, Germany

María Espinosa and Sergio Gomez y Paloma
European Commission, Agriculture and Life Sciences in the Economy, Sevilla, Spain

Abstract

The prediction of farm structural change is of large interest at EU policy level, but available methods are limited regarding the joint and consistent use of available data sources. This paper develops a Bayesian Markov framework for short-term prediction of farm numbers that allows combining two asynchronous data sources in a single estimation. Specifically, the approach allows combining aggregated Farm Structure Survey (FSS) macro data, available every two to three years, with individual farm level Farm Accountancy Data Network (FADN) micro data, available on a yearly basis. A Bayesian predictive distribution is derived from which point predictions such as mean and other moments are obtained. The proposed approach is evaluated in an out-of-sample prediction exercise of farm numbers in German regions and compared to linear and geometric prediction as well as a "no-change" prediction of farm numbers. Results show that the proposed approach outperforms the geometric prediction but does not significantly improve upon the linear prediction and a prediction of no change in this context.

Key Words

Bayesian prediction; Markov transitions; asynchronous data; structural change

Zusammenfassung

Die Vorhersage des landwirtschaftlichen Strukturwandels ist von großem Interesse für die EU-Agrarpolitik, aber gegenwärtige Methoden können die verfügbaren Datenquellen nicht vollständig und in konsistenter Weise verwenden. Zur kurzfristigen Vorhersage des Agrarstrukturwandels wird ein Bayes'scher Markov-Ansatz entwickelt, der die Kombination von zwei asynchronen Datenquellen in einer einzigen Schätzung erlaubt. Im konkreten Fall werden dabei in konsistenter Weise aggregierte Daten des Farm Structure Survey (FSS), die alle zwei bis drei Jahre erhoben werden, mit jährlich verfügbaren Stichprobendaten des einzelbetrieblichen Farm Accountancy Data Network (FADN) kombiniert. Eine geschätzte Bayes'sche Vorhersageverteilung erlaubt die Ermittlung von Punktvorhersagen in Form des arithmetischen Mittels und die Ableitung anderer Momente. Evaluiert wird der Ansatz in einer „out-of-sample"-Vorhersage für die Anzahl landwirtschaftlicher Betriebe in verschiedenen Klassen und Bundesländern in Deutschland. Verglichen werden die Ergebnisse mit einer linearen und geometrischen Vorhersage sowie einer „konstanten" Vorhersage, die keine Veränderungen zum letzten Beobachtungsjahr unterstellt. Im Vergleich zur geometrischen Vorhersage liefert der Ansatz bessere Ergebnisse, wobei lineare und konstante Vorhersage ähnliche Ergebnisse in diesem Kontext liefern.

Schlüsselwörter

Bayes'sche Vorhersage; Markov-Übergänge; asynchrone Daten; Strukturwandel

1 Introduction

Detailed up-to-date information about farm structural change is of great interest for policy makers and stakeholders and provides the basis for policy analysis. Farm structural change relates to medium to long-term investment decisions to enter or leave the business or to fundamentally change farm size, specialisation or production intensity. Therefore, modelling of farm structural adjustments in an ex-ante policy modelling exercise is highly relevant (ZIMMERMANN et al., 2009). The Common Agricultural Policy post 2013 (EPEC, 2013) has introduced a new design of direct payments that may have important implications for farm structural change adjustments. In particular, the new basic payment scheme is no longer based on uneven historical references, but rather on converging per ha payment at national or regional level. The change in decoupled support (pre- and post-2013 reform) may affect farmers' decision related to entry/exit into the sector (MORO and SCKOKAI, 2013). Furthermore, member states have the option to further adjust direct payments through a redistributive payment to the first hectares of farms. This payment provides more targeted support to small and medium-sized farms, therefore creating incentives for reducing farm size. In addition, member states may also grant limited coupled support to specific vulnerable crops, therefore changing the profitability of specific crops with a potential effect on crop composition (i.e. change in farm specialisation). However, the overall assessment of the impact of the CAP post-2013 on farm structural adjustments will depend on the final implementation by the member states. Nevertheless, most of the policy assessments developed at farm and farm-type levels are limited to the first order effect, and therefore do not take into account the adaptation of individual farms/farm-types to market and policy changes. Consequently, upscaling the results of the models developed at farm/farm-type require a prediction of the number of farms (and its aggregation in each farm-type) for future scenarios.

Apart from this more general relevance of predicting farm structural change in the context of upscaling results of farm models to regional level, there is also a specific context of the research presented in this paper. The "Economic Analysis of EU Agriculture" Unit of DG-Agri of the EU-commission regularly needs to predict farm numbers in farm size and specialisation classes in between the years of the Farm Structure Survey (FSS). The underlying research is funded in the context of a project aiming to improve the methodology for providing such rather short-term forecasts by using transitions observed in the sample of the Farm Accountancy Data Network (FADN) (GOCHT et al., 2013).

Methodologically, this paper contributes to the literature in three ways: 1) it allows to consistently combine the bi- or triennial FSS data with the yearly FADN data in the estimation of yearly transition probabilities (TPs), thereby improving upon previous approaches with data interpolation as in ZIMMERMANN and HECKELEI (2012b). The approach developed in STORM et al. (2015) allows merging such asynchronous data sources in a single estimation explicitly reflecting their connection in the data generating process; 2) in contrast to STORM et al. (2015) a Parallel Tempering (PT) approach (LIU, 2008) for sampling from the posterior is implemented replacing the simple Metropolis-Hasting sampler. The PT sampler converges more reliably when faced with a multimodal posterior distribution; 3) we develop a Bayesian prediction framework that offers a predictive distribution for the number of farms in classes from which point predictions and predictive uncertainty can be derived.

The approach is illustrated and evaluated in an out-of-sample prediction for seven (West) German Regions for which a relatively long sample is available. We predict farm numbers for different size classes, with and without differentiation of specialisation classes. Specifically, three (economic) size classes and an entry/exit class are considered for different aggregation levels regarding farm specialisation. First, we perform a prediction at an aggregated level where farm numbers in different size classes and the entry/exit class are predicted without any distinction by farm specialisation. Then the prediction for the three size classes and entry/exit is repeated at a more disaggregated level for three different farm specialisations, namely crop, livestock and mixed farms. In each case, three different time periods are considered in the out-of-sample prediction. The predictions based on the Markov approach are compared to simple constant, linear and geometric predictions of farm numbers.

Even though we choose seven West Germen regions for illustrative purposes, it should be pointed out that the approach can be directly transferred to other EU member states for sufficiently long series of FSS and FADN, currently available in at least the EU-15 member states.

The remaining structure of the paper is as follows: the following section describes the data sources. Section 3 develops the estimation and prediction framework and derives an appropriate measure to assess the performance of the Bayesian Markov approach compared to its simple alternatives. Section 4 discusses the specific implementation, including the setup of the out-of-sample prediction, the selection of explanatory variables and the implementation of the PT sampling algorithm. Section 5 presents the results and section 6 offers some concluding remarks.

2 Data

FSS and FADN are the two major data sources suitable for the analysis of farm structural change at EU level, both providing information at farm level for all EU member states. In this paper we aim to combine both data sources for a more precise prediction of farm structural change. The developed approach allows completing information on farm numbers in size and specialisation classes between FSS years, and predicting these variables in a short- or medium-term time horizon. The focus here is in short-term predictions (2-4 years).

The FSS is a census of all agricultural holdings conducted every ten years with three intermediate surveys (census or sample surveys depending on each MS) conducted in-between (Council Regulation (EC) No 1166/2008). FSS data is thus available every two to three years offering aggregated information about the total number of farm holdings in different size or specialisation classes[1]. In the following, we refer to this aggregated data as *macro data*. On the other hand, FADN data is available on a yearly basis and provides information about class transitions of individual farms for a sample of commercial farms. Table 1 presents the FADN and FSS data availability at the time when this study was conducted. Different from FSS, FADN data allow tracking the development of one farm in the sample over several years including the movement of farms between classes. We will refer to this type of data in the following as *micro data*. The sample of FADN farms shall represent all relevant farm types and farm sizes in each region according to a stratified sampling plan[2].

Table 1. Available FADN and FSS years

Year	FADN years (t)	FSS years (λ)
1989	0	
1990	1	0*
1991	2	
1992	3	
1993	4	1
1994	5	
1995	6	2
1996	7	
1997	8	3
1998	9	
1999	10	
2000	11	4*
2001	12	
2002	13	
2003	14	5
2004	15	
2005	16	6
2006	17	
2007	18	7
2008	19	

*Years where FSS is a full census, in other years information is derived based on a sample survey.
Source: FADN and FSS data base

Given the shorter intervals with which FADN data is collected and the shorter release time, FADN data is generally the more recent information on farm numbers in classes compared to FSS. Therefore, we might have FADN data for up to three more years after the last available FSS year. This paper exploits this information – together with all other available FADN micro and FSS macro data from previous years – to predict farm numbers in different size classes (including an entry/exit category).

3 Methodology

3.1 Bayesian Estimation Framework

Following STORM et al. (2015) the number of farms in different classes is modelled as a Markov process. In a Markov process, the movement of individuals between a finite number of predefined, mutually exclusive, and exhaustive states, $i = 1,...,k$, is a stochastic process. In the following we consider a situation in which the states represent an entry/exit and three different farm size classes ($k = 4$). The Markov process is characterized by a $(k \times k)$ TP matrix $\mathbf{P}_t$ (in the following bold letter denote matrices or vectors). The elements P_{ijt} of that matrix give the probability that an individual moves from state i in $t-1$ to j in t. The

[1] The individual level (micro) FSS data is processed by the individual member states and typically not accessible for confidentiality reasons, whereas FSS macro data is publically available.

[2] For each sample farm, however, a weight is calculated using the information in FSS about the total number of farms in each farm type, size class and region. With these weights, the FADN sample can be aggregated to match FSS results on the population level and information about the total number of farms in each farm type or size class (macro data) can be derived. The weights, however, still reflect only the last available FSS year.

$(k \times 1)$ vector $\mathbf{n}_t$ denotes the number of individuals in each state i and develops over time according to a first order Markov process.

$$\mathbf{n}_t = \mathbf{P}_t' \mathbf{n}_{t-1}. \tag{1}$$

In a non-stationary[3] Markov process the TPs change over time depending on exogenous variables. The way the exogenous variables relate to the TPs, $\mathbf{P}(\boldsymbol{\beta})$ differs depending on the type of Markov states considered. If we assume that the Markov states do not have an order, the relationship between exogenous variables and TPs should be specified based on the multinomial logit model, whereas an ordered logit model is suitable for our case where transitions between size classes are considered (see STORM et al., 2015).

For the estimation of the non-stationary TPs a Bayesian estimation framework is employed that allows combining macro and micro data in the estimation of non-stationary Markov TPs. For a detailed description we refer to STORM et al. (2015). The general idea of the framework is combining a macro data based likelihood function with a micro data based prior density. Both likelihood and prior are therefore data based and represent the two different available data sources in a consistent manner. Similarly as in STORM et al. (2015) we will combine FSS macro data, available every two to three years, with the FADN micro data, available at a yearly base.

The prior density is combined with the likelihood function to a posterior density used for deriving the marginal density of individual parameters. Since the required integration is not traceable analytically, we employ a Monte Carlo Integration approach. STORM et al. (2015) use a simple Metropolis-Hastings (MH) algorithm to draw a sample from the posterior. Here we replace the MH algorithm by a Parallel Tempering (PT) sampling algorithm (LIU, 2008). The general idea of the PT approach is to run multiple copies of the original chain raised to different powers (in the PT context called "temperatures") in parallel and allow exchanges between them. The advantage of the PT approach is that the 'heated' chains (raised to powers smaller than one) are able to escape local modes more easily such that it becomes easier to sample from multimodal posterior distributions like those found in the specific application.

[3] Note that non-stationary in the Markov context describes the fact that the transition probabilities are allowed to differ over time depending on explanatory variables.

In our particular case, we adopt the following setup of the PT sampler. We consider I parallel chains with temperatures $1 = T_1 < T_2 < \ldots < T_I$. The PT sampler consists of parallel and swapping steps. In each parallel step r the current states $x_1^{(r)}, x_2^{(r)}, \ldots, x_I^{(r)}$ of all I chains are updated in simple MH steps using a random walk MH sample with a multivariate normal proposal density. After every five parallel steps a swap between all neighbouring chains is proposed. Denoting neighbouring chains as i and $i+1$, a swap of states $x_i^{(r)}$ and $x_{i+1}^{(r)}$ is accepted with probability

$$\min\left\{1, \exp\left(\pi\left(x_i^{(r)}\right) - \pi\left(x_{i+1}^{(r)}\right)\right)^{T_i - T_{i+1}}\right\}, \tag{2}$$

where $\pi\left(x_i^{(r)}\right)$ denotes the log posterior density evaluated at state $x_i^{(r)}$ of chain i. Swaps are first considered for the last two chains and then going back in steps to the first two neighbouring chains. With such a setup it is generally possible that the state of the last chain, $x_I^{(r)}$, is swapped to the first chains within one pass through all neighbouring chains. This setup was found to be more efficient in our specific application compared to the approach proposed by LIU (2008) in which only one pair of neighbours are selected at random to swap states in each step.

The performance of the PT tempering crucially depends on the chosen number of parallel chains, I, as well as on the chosen temperatures $1 = T_1 < T_2 < \ldots < T_I$ and the covariance matrices of the multivariate normal proposal densities to be selected for each specific sampling. The temperatures require covering a sufficiently large temperature range such that the hottest chain can easily escape local modes. On the other hand, the differences between neighbouring chains' temperatures need to be small enough such that a sufficient amount of swaps are accepted. The specific implementation of the PT approach is described in section 4.3.

3.2 Prediction Methods

The Markov process specified in (1) may be directly used for prediction of farm numbers in different states. The number of farms in k states in the last observed year t is denoted by a $(k \times 1)$ vector $\mathbf{n}_t$. Our aim is to predict farm numbers $\hat{\mathbf{N}} = \hat{\mathbf{n}}_{t+1}, \ldots, \hat{\mathbf{n}}_{t+\tau}$ in k states for τ years starting from the last observed year t. Taken the TPs $\mathbf{P} = (\mathbf{P}_{t+1}, \ldots, \mathbf{P}_{t+\tau})$ as given, prediction to $t+\tau$ follows directly from (1) by

$$\hat{\mathbf{n}}_{t+\tau} = \left(\prod_{v=t+1}^{t+\tau} \mathbf{P}_v \right)' \mathbf{n}_t \,. \qquad (3)$$

With (3) the predicted farm number $\hat{\mathbf{N}}$ are thus a function of the TPs, $\mathbf{P}$. The TPs are itself a function of the unknown parameter $\boldsymbol{\beta}$, thus we can write $\hat{\mathbf{N}} = \hat{\mathbf{N}}(\mathbf{P}(\boldsymbol{\beta})) = \hat{\mathbf{N}}(\boldsymbol{\beta})$. The specification of the functional relationship $\mathbf{P}(\boldsymbol{\beta})$ is based the ordered logit specification (see STORM et al., 2015).

The Bayesian estimation framework provides several ways of how to implement the prediction. One possibility is to derive point estimates of $\boldsymbol{\beta}$ such as the posterior mean, which is the optimal Bayesian estimator under squared error loss. Here we employ an alternative prediction strategy directly using the sample outcomes of the joint posterior of $\boldsymbol{\beta}$. This provides the advantage that a complete Bayesian predictive distribution is derived for each state and year in an intuitive and straightforward way. Technically, each sample outcome $\boldsymbol{\beta}_{(l)}$, $l = 1,...,L$ from the posterior is used to predict farm numbers based on (3) obtaining a sample of predictions $\hat{\mathbf{N}}_{(l)} = \hat{\mathbf{N}}(\boldsymbol{\beta}_{(l)})$. This sample can be regarded as a sample from the predictive distribution $f\left(\hat{\mathbf{N}} \middle| \mathbf{d}\right) = \hat{\mathbf{N}}(\boldsymbol{\beta}_{(l)})h(\boldsymbol{\beta}|\mathbf{d})$. The predictive distribution may itself be the final result or alternatively summary statistics like mean, variance and quintiles of the predictive distribution may be provided from the obtained sample.

3.3 Prediction Measures

The prediction quality of the described approach is compared to the simple linear, constant and geometric prediction based on the Mean Absolute Scaled Error (MASE). The MASE is proposed by HYNDMAN and KOEHLER (2006) who argue that the MASE is superior to other commonly used forecast measures such as the (Root) Mean Square Error (which is not scale free); measures based on relative errors, such as the Mean Relative Absolute Error, or relatives measures, such as the relative Mean Absolute Error. The MASE has a clear interpretation, is scale free and is defined in all relevant situations (only in the irrelevant case where historical data shows no variation it is not defined). It is calculated by dividing the absolute prediction error $e_t = \left|\hat{Y}_t - Y_t\right|$, where $\hat{Y}_t$ is a prediction of Y_t, by the average one-step naive forecast in the sample period,

$$MASE = mean \frac{e_i}{\frac{1}{n-1}\sum_{i=2}^{n}\left|Y_t - Y_{t-1}\right|} \,. \qquad (4)$$

Therefore, a MASE less than one indicates a better prediction than the average one-step naive forecast within-sample. In our specific case the MASE is calculated for the predictions of farm numbers in $t+\tau$ over all regions and size classes, without considering the artificial entry/exit class. The average one-step naive forecast is calculated over all observed FSS years. This is relevant for the interpretation of the absolute size of the MASE since the step-length of the out-of-sample prediction might differ from the step length of the naive one-step forecast (two or three years). It is, however, irrelevant for a relative comparison of the MASE between different prediction methods being the primary purpose of the out-of-sample prediction.

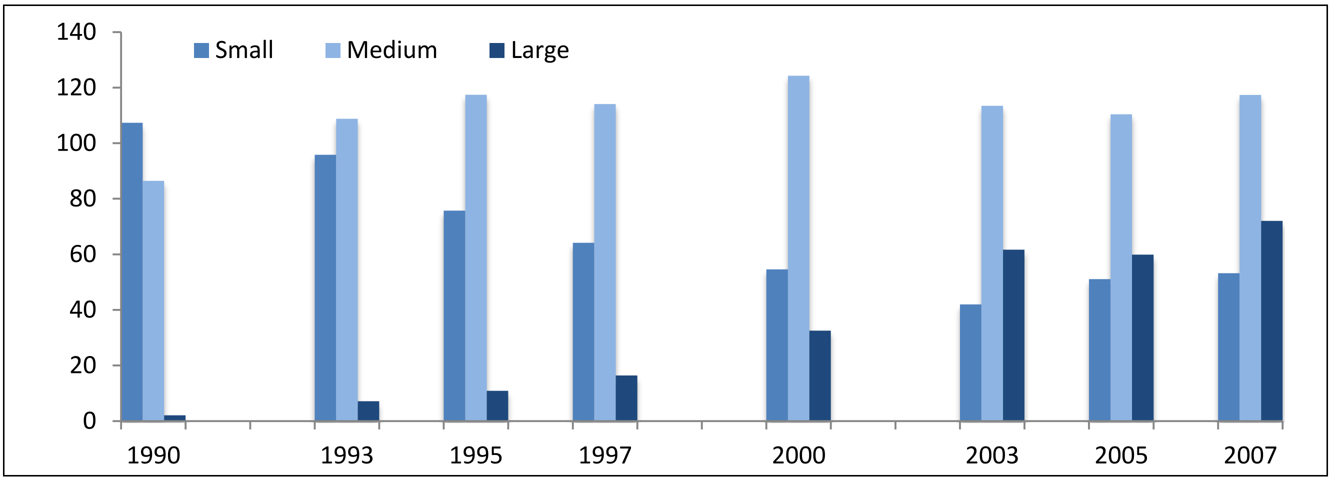
4 Implementation

4.1 Setup of Out-of-Sample Prediction

In the out-of-sample prediction, farm numbers are predicted for different size and specialisation classes. The classification of farms is based (in FADN and FSS) on economic size and specialisation classes (Commission Decision 85/377/EEC). The physical units of production (hectare or livestock units) are valued by the corresponding Standard Gross Margins (SGM) calculated for each region on a regular basis by the member states. The sum of all production activities valued by the SGM determines the economic size of a farm, expressed in Economic Size Units (ESU), while the share of each production activity on total ESU determines the farm specialisation.[4]

[4] From the accounting year 2010, the typology for agricultural holdings is based on Standard Output (Commission Regulation (EC) No. 1242/2008) instead of SGM. The main differences among the SGM and SO is that the SO excludes direct payments and the cost of variable inputs. Moreover, the unit used to measure SO is the Euro and not the Economic Size Unit (1.200 Euro). The change will have no effect on the general applicability of the proposed prediction approach.

Figure 1. Distribution of farm numbers within the three size classes for each FSS year



Note: regions: Schleswig-Holstein (FADN code: 10), Lower Saxony (30), North Rhine-Westphalia (50), Hesse (60), Rhineland-Palatinate (70), Baden-Württemberg (80) and Bavaria (90). Farm types: TF14: 13,14, 41,42,43,44,45,50,60,70,71,72,80,81,82.
Source: own illustration based on FADN data

In the out-of-sample prediction, four different situations are distinguished. On the one hand the prediction is performed for all farms (excluding horticulture and permanent crops *TF14*: 20, 31, 32, 33, 34) irrespectively of their farm type. Additionally, the prediction is repeated for three different farm specialisations, namely crop farms (*TF14*: 12, 14, 60), livestock farms (*TF14*: 41, 44, 45, 50, 70) and mixed farms (*TF14*: 80). In principle this rather broad classification in three types can be extended but is limited by the increasing number of zero observations resulting from the increasing number of combinations of farm type, region, year and size class that arise when further differentiating farm types. In each of the four cases, three different size classes (small 16-40ESU, medium 40-100ESU and large >100ESU) and an entry/exit class are considered. The definition of the size classes is predefined by the FADN data and the only three size classes that can consistently used from 1990-2007[5]. Figure 1 illustrates the development of the distribution of farm number in the three size classes. The entry/exit class is an artificial class required by the Markov approach and representing farms that enter or quit farming (STOKES, 2006). It should be pointed out that FADN considers only "commercial farms" defined as those farms above 16ESU. All farms below that threshold are considered "hobby" farms in FADN and there are treated as farm exits in our analysis. Farm structural change that happens below this threshold remains unobserved. This needs to be kept in mind we interpreting the results.

[5] For Germany FADN considered a size class from 8-16ESU until 1997 and since 1999 FADN differentiated the "large" size class (>100ESU) in 100-250ESU and >250ESU.

Table 2. Out-of-sample prediction periods and corresponding data considered for estimation

Prediction period	FADN data considered in estimation	FSS data considered in estimation
2000-2003	1989-2003	1989-2000
2003-2005	1989-2005	1989-2003
2005-2007	1989-2007	1989-2005

Source: own table

For each of the four cases, three different out-of-sample prediction periods are considered (Table 2). In each prediction period the last FSS year is excluded from estimation and macro data instead predicted for that year. The prediction is then compared to the observed macro data. By considering three different time periods, each time excluding an additional FSS year in estimation, it can be evaluated how the approach would have performed in previous periods. Table 2 presents the different prediction periods and the corresponding FADN and FSS data used.

For each individual prediction, a panel of seven West-German regions is considered in estimation (FADN regional codes: 10 (Schleswig-Holstein), 30 (Lower Saxony, 50 (North Rhine-Westphalia), 60 (Hesse), 70 (Rhineland-Palatinate), 80 (Baden-Württemberg) and 90 (Bavaria)).

These 12 different Bayesian Markov predictions (three time periods for each of the four cases) are compared to a constant, linear and geometric prediction. The linear prediction employs a least squares estimation of γ_1 and γ_2 of the linear function $n_t = \gamma_1 + \gamma_2 t + \varepsilon_t$, where n_t is the number of farms in time t. Using the estimates $\hat{\gamma}_1$ and $\hat{\gamma}_2$, farm numbers for $t+1$ are then predicted by $\hat{n}_{t+1} = \hat{\gamma}_1 + \hat{\gamma}_2(t+1)$ and for the following years accordingly. For the estimation only FSS macro data is employed. The geometric growth rate is derived by a least squares estimation of $\ln(n_t) = \lambda_1 + \lambda_2 t + \varepsilon_t$. Farm numbers in $t+1$ are predicted using the estimated parameters $\hat{\lambda}_1$ and $\hat{\lambda}_2$ to calculate $\hat{n}_{t+1} = e^{(\hat{\lambda}_1 + \hat{\lambda}_2(t+1))}$. Data source and time periods are the same as those used for the linear prediction. An advantage of the geometric over the linear prediction is that predicted farm numbers cannot become negative. Problems arise, however, in the geometric prediction in cases in which no farms are observed in a particular time period. In these cases, the dependent variable is not defined, and we omit the observation from the estimation. The constant prediction assumes that farm numbers do not change during the prediction period, such that the predicted value is equal to the last observed value for each farm type and region.

4.2 Identification of Potential Explanatory Variables

To select a set of explanatory variables for the estimation of the non-stationary TPs, first a set of factors that potentially drive farm structural change are identified based on theoretical considerations and the literature analysing factors influencing farm structural change (BREUSTEDT and GLAUBEN, 2007; ZIMMERMANN et al., 2009; PIET et al., 2012; ZIMMERMANN and HECKELEI, 2012a; ZIMMERMANN and HECKELEI, 2012b). The identified factors may broadly be categorized in six general categories: technology, the initial farm structure, market conditions, natural resource factors, social and demographical factors and agriculture policy (see Table 3). For each potential factor, specific explanatory variables are identified that allow approximating that factor.

The model is specified as a fixed effects model with regional dummy variables included for each region except one. These dummy variables capture all time invariant factors such as the initial farm structure (farm size, size heterogeneity), natural conditions (share of absolute grassland, slope, climate, population density etc.) that remain rather stable over the data period. Clearly, some of these factors also vary substantially within regions (e.g. slope, population density) which are relevant for individual farms but not captured in a regional level study. For off-farm

Table 3. Factors identified to potentially influence farm structural change and corresponding explanatory variables

General Category	Factors	Approximated by
Technology	Yields	Index of Standard Gross Margins (SGM) for different farm specialisations. Specialist COP (*SGM13*), Specialist other filed crops (SGM14), Specialist Milk (SGM41), Specialist sheep/goats/cattle (SGMLive), Specialist Grainivores (SGM50) *Source: FADN*
Initial farm structures	Farm size/capacity	*Captured in fixed effects*
	Size heterogeneity	*Captured in fixed effects*
Market conditions	Input/output prices (price ratios)	SGMs (see Technology)
Natural resource factors	Share of grassland	*Captured in fixed effects*
	Slope	*Captured in fixed effects*
	Climate	*Captured in fixed effects*
Social and demographical factors	Population density/growth	*Captured in fixed effects*
	Off-farm income opportunities	Unemployment rate (*Unemp*) *Source: DeStatis*
	Age structure	Percentage of farmers aged above 60 (*Above60*) *Source: FADN*
Agricultural Policy	Agricultural Policy	Dummy variables for major policy reforms (MacSharry reform, Agenda 2000 and Midterm review)

Source: BREUSTEDT and GLAUBEN (2007); ZIMMERMANN et al. (2009); PIET et al. (2012); ZIMMERMANN and HECKELEI (2012a); ZIMMERMANN and HECKELEI (2012b) employment opportunities, the unemployment rate and for the age structure of the farm population the percentage of farmers above 60 years old are considered as explanatory variables. Agricultural policy is considered by three dummy variables indicating major shifts in EU Agricultural Policy in 1993 (MacSharry Reform), 2000 (Agenda 2000) and 2003 (Mid-Term Review).

Technological developments as well as market conditions are represented by standard gross margins (SGM) for different production activities as explanatory variables. SGMs are provided by EuroStat (Commission Decision 85/377/EEC) at regional level for all relevant production activities and member states. SGMs are calculated by member states based on a period of several years to reduce the effects of short-term price or yield fluctuations. Therefore, SGMs should reflect longer-term changes in productivity as well as in input or output prices that affect the attractiveness of different production activities. For our purpose, we aggregate the different individual SGMs into five SGM indices reflecting major production specialisation activities. Specifically, SGMs indices are calculated for *Specialist COP (SGM13)*, *Specialist other filed crops (SGM14)*, *Specialist Milk (SGM41)*, *Specialist sheep/goats/cattle (SGM40)* and *Specialist Granivores (SGM50)*. It is assumed that the SGMs affect transitions of farms between classes in two different ways. On the one hand, SGMs reflect the productivity of production factors in different activities. Hence an increase of the SGM of one specialisation should increase the attractiveness of the corresponding farm type. This in turn draws production factors and therefore farms into those farm specialisations. Also, an increase in the SGM should lead to an increase of the ratio of on-farm to off-farm income possibilities, such that farm entries/exits should increase/decrease. On the other hand, changes in SGMs directly affect the transitions between states because the classification of farms in size classes depends on the SGMs. Therefore, changes in SGMs also have a direct effect on the change between classes. An increase in SGM increases the economic size of a farm even though the physical layout stays the same; hence the farm would move to a higher size class. These two effects, movements in the physical units as well as in the valuation of each unit, render an interpretation of the causal relationship between SGM and farm structural change somewhat difficult but this is irrelevant for the prediction of farm numbers.

The set of explanatory variables is further restricted using the high correlation between individual explanatory variables. Particularly, three SGM indices (*Specialist other filed crops (SGM14)*, *Specialist sheep/goats/cattle (SGM40)* and *Specialist Granivores (SGM50))* are excluded because they are highly correlated to the other two SGMs (Table 4). Even though high correlations among explanatory variables are irrelevant for prediction they add little to the overall explanatory power of the model, and are therefore excluded in order to limit the numerical complexity which increases with each additional explanatory variable.

The final set of explanatory variables use in estimation thus consists of the regional fixed effects, the SGM for *Specialist COP (SGM13)* and *Specialist Milk (SGM41)* the unemployment rate (*Unemp*), the percentage of farmers above 60 years old (*Above60*) the three dummy variables for the MacSharry Reform, Agenda 2000 and the Mid-Term Review. Obviously, this set of explanatory variables is clearly limited to explain the complex processes driving structural decisions by individual farmers. However, the choice is motivated by the aim to find variables that are generally available for all EU regions. Also, the rather broad classification of farms in three size classes and three farm types as well as the regional focus of the study limit the identification of relevant explanatory variables. The farm population within one class or

Table 4. Correlation matrix of explanatory variables

	SGM13	SGM14	SGM41	SGM40	SGM50	Unemp	Above60
SGM13	1	**0.84**	-0.13	0.25	0.10	0.41	0.36
SGM14		1	0.11	0.49	0.34	0.40	0.40
SGM41			1	**0.85**	**0.82**	0.06	0.10
SGM40				1	**0.92**	0.19	0.35
SGM50					1	0.19	0.23
Unemp						1	-0.16
Above60							1

Source: own calculation

Table 5. Mean, 5% and 95% Quintiles of the marginal posterior density for the first 10 of 64 coefficients estimated in five identical runs using different random starting values

Coef.	1. Run	2. Run	3. Run	4. Run	5. Run
Mean of the marginal posterior density					
1	-3.08	-3.01	-3.08	-3.14	-3.01
2	0.86	0.86	0.87	0.86	0.85
3	3.81	3.75	3.77	3.78	3.79
4	-2.93	-2.89	-2.88	-2.95	-2.89
5	1.02	1.02	1.02	1.03	1.01
6	-0.58	-0.59	-0.59	-0.58	-0.60
7	-0.64	-0.65	-0.64	-0.64	-0.64
8	0.88	0.90	0.89	0.87	0.89
9	1.81	1.75	1.80	1.84	1.76
10	1.58	1.57	1.57	1.58	1.58
5% Quintiles of the marginal posterior density					
1	-3.43	-3.52	-3.49	-3.52	-3.46
2	0.69	0.62	0.71	0.71	0.08
3	3.13	3.37	3.18	3.00	3.28
4	-3.17	-3.14	-3.16	-3.16	-3.16
5	0.73	-0.15	0.54	0.64	-0.06
6	-0.81	-0.74	-0.74	-0.72	-0.85
7	-0.78	-0.81	-0.78	-0.81	-0.77
8	0.60	0.69	0.69	0.67	0.53
9	1.25	0.54	1.43	1.44	0.83
10	1.37	1.42	1.41	1.25	1.45
95% Quintiles of the marginal posterior density					
1	-2.18	-1.13	-2.43	-2.41	-0.13
2	1.19	1.08	1.15	1.33	1.01
3	4.00	3.95	3.92	3.93	4.01
4	-1.11	-1.32	-1.30	-1.00	-0.08
5	1.34	1.31	1.25	1.22	1.38
6	-0.45	-0.15	-0.47	-0.43	-0.14
7	-0.45	-0.47	-0.49	-0.52	-0.42
8	1.17	1.19	1.17	1.20	1.16
9	2.01	2.03	2.03	2.06	2.01
10	1.71	1.82	1.71	1.71	2.09

Note: estimation is for the prediction of crop, livestock and mixed farms combined for the prediction period from 2005 to 2007.
Source: own estimation

region is likely to exhibit a substantial degree of heterogeneity such that specific explanatory variables might have different effects on individual farms within one class. For example, the estimation for "livestock" farms groups together specialized milk farms with extensive sheep farms that might react vary differently to specific explanatory variables. However, as discussed above the possibilities to consider a more detailed resolution with respect to size classes or farm types is limited due to data constraints.

4.3 Implementation of the Parallel Tempering Sampler

For sampling from the posterior we found an implementation of the PT approach using $I = 30$ parallel chains to be suitable for delivering robust sample results. The selection of temperatures and covariance matrixes of the proposal densities requires a substantial amount of manual fine tuning for each individual estimation. Temperatures are chosen such that the swap acceptance rate is above 20% for most of the pairs and at least 2-3% such that swaps between all chains are possible. The covariance matrices of the multivariate normal proposal densities are specified as diagonal matrices with equal variance for all parameters within one chain but differ across chains such that an acceptance rate between 20-30% is obtained for most chains. Starting values for all parameters in all chains are drawn randomly from a uniform distribution with the specific support chosen for the parameter. For the final estimation a burn-in period of two mill. draws and a sample of one mill. draws are used. Computations are performed using Aptech's GAUSS™ 12 on an Intel® Xeon® E5-2690, where computation time for one estimation is around 1.6 hours using around half of the available CPU.

5 Out-of-Sample Prediction Results

To assess the quality of the different prediction approaches different measures based on the Absolute Scaled Error are considered. In the out-of-sample prediction we obtain prediction results in each of the four cases for seven regions, three time periods and three size classes.[6] For each single prediction, the Absolute Scaled Error is calculated and then summarized across predictions by the mean and median Absolute Scaled Error as a measure of central tendency as well as the standard

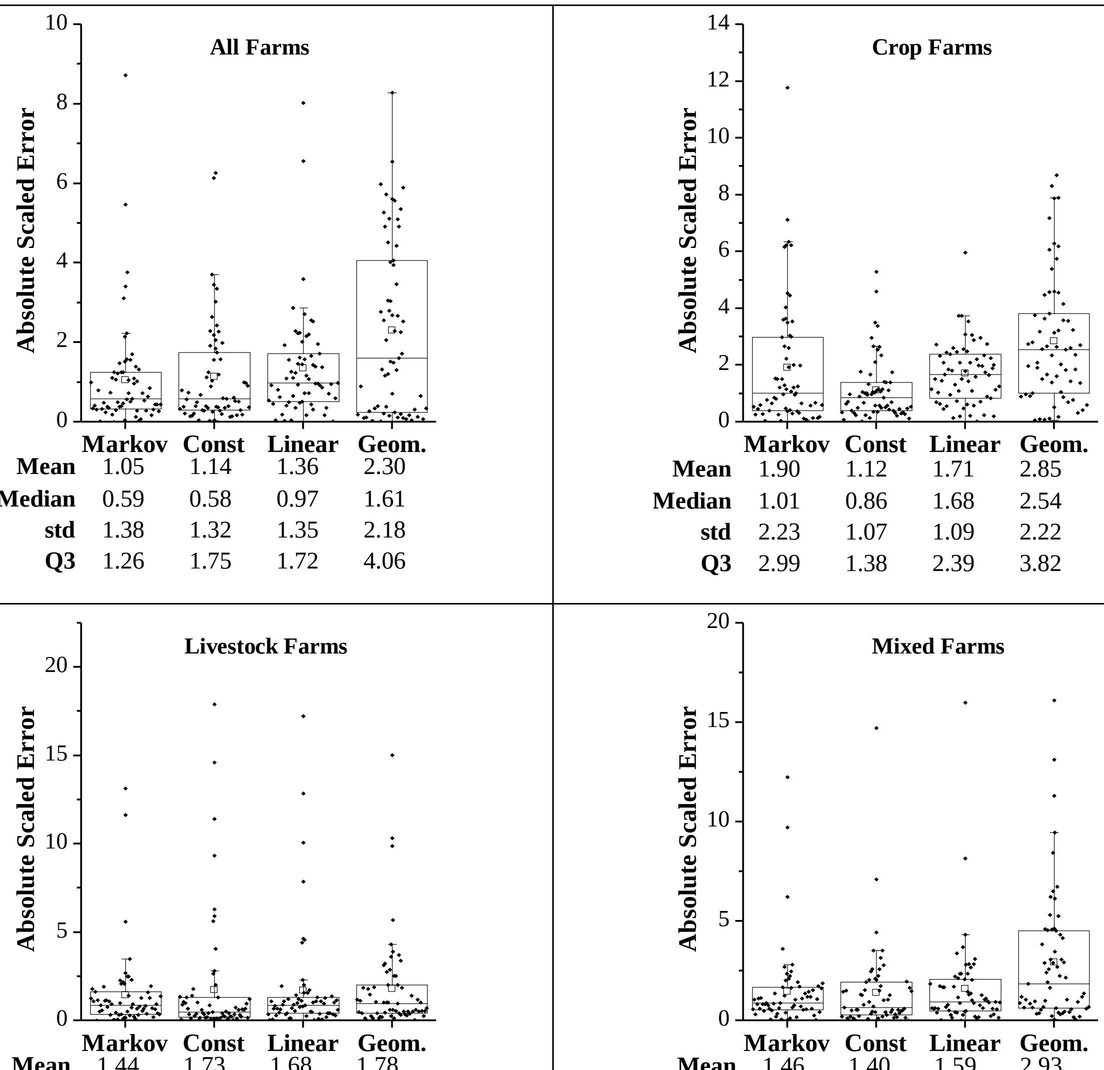
[6] The prediction for the entry/exit class is not considered since it is a not observed artificial class (see section 3.1).

deviation and the 3rd quartile as measures of spread. The 3rd quartile is used as we are only interested in how far the Absolute Scaled Error deviates from zero.

Figure 2 depicts the performance measures of the different prediction method for the four different cases considered. The Markov approach clearly outperforms the geometric prediction in all four cases with respect to all measures. Compared to the linear prediction and the prediction of no change the picture is less clear. With respect to the mean Absolute Scaled Error, the Markov prediction outperforms the constant and the linear prediction in case of 'all' farms and livestock farms while it is outperformed by the constant and linear prediction in case of crop farms and the constant prediction in case of mixed farms. With respect to the median Absolute Scaled Error the Markov prediction is slightly inferior to the prediction of no change which has either a very similar or slightly lower median Absolute Scaled Error. Compared to the linear prediction, the Markov prediction is superior except for the case of Livestock farms where the linear prediction is slightly better.

For an overall assessment the individual results of the four cases are combined to obtain an overall measure of the prediction quality. The results are given in Figure 3. The geometric prediction performs worst on all measures, followed by the linear prediction. The Markov prediction is slightly better than the linear prediction and slightly worse than the constant prediction with respect to the mean, median and

Figure 2. Box-Whisker-Plot of Absolute Scaled Errors for different prediction methods in four cases



All Farms

	Markov	Const	Linear	Geom.
Mean	1.05	1.14	1.36	2.30
Median	0.59	0.58	0.97	1.61
std	1.38	1.32	1.35	2.18
Q3	1.26	1.75	1.72	4.06

Crop Farms

	Markov	Const	Linear	Geom.
Mean	1.90	1.12	1.71	2.85
Median	1.01	0.86	1.68	2.54
std	2.23	1.07	1.09	2.22
Q3	2.99	1.38	2.39	3.82

Livestock Farms

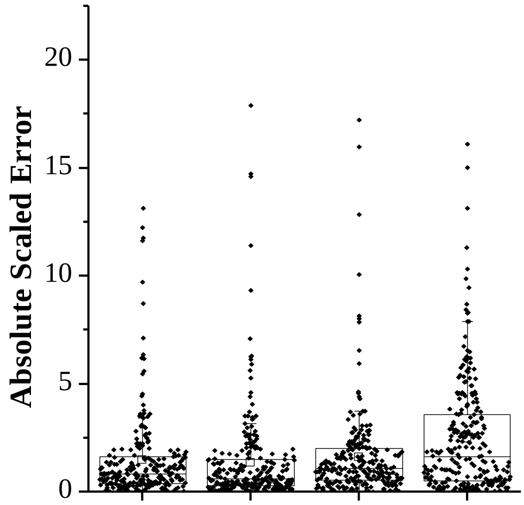
	Markov	Const	Linear	Geom.
Mean	1.44	1.73	1.68	1.78
Median	0.87	0.48	0.86	0.94
std	2.23	3.42	3.00	2.62
Q3	1.62	1.31	1.33	2.03

Mixed Farms

	Markov	Const	Linear	Geom.
Mean	1.46	1.40	1.59	2.93
Median	0.87	0.65	0.94	1.84
std	2.02	2.13	2.27	3.28
Q3	1.68	1.92	2.08	4.53

Note: absolute Scaled Errors are displayed for each prediction in three size classes, three prediction periods and seven regions considered. In each case the Markov prediction is calculated as the mean of the posterior predictive distribution.
Source: own estimation

Figure 3. Box-Whisker-Plot of Absolute Scaled Errors for different prediction methods in the out-of-sample prediction



	Markov	Const	Linear	Geom.
Mean	1.46	1.35	1.59	2.46
Median	0.85	0.62	1.08	1.64
std	2.00	2.19	2.07	2.64
Q3	1.61	1.52	2.00	3.59

Note: Absolute Scaled Errors are displayed for a prediction of crop, livestock and mix farms as well as for a prediction of all farms combined. In each case farm number are predicted in three size classes, three prediction periods and seven regions. The Markov prediction is calculated as the mean of the posterior predictive distribution.
Source: own estimation

3rd quartile of the Absolute Scaled Error. The Markov prediction has a slightly lower standard deviation compared to the linear and constant prediction. The results indicate that overall the Markov prediction is not able to clearly outperform the prediction of no change. The strong performance of the constant prediction in comparison to the Markov prediction but particularly in comparison to the linear prediction indicates that the farm structure is rather stable within short intervals of two to three years and does not follow a clear trend.

6 Conclusion

Overall, the paper contributes to the literature by extending the Bayesian estimation approach for non-stationary Markov model developed in STORM et al. (2015) by implementing a Parallel Tempering sampler that allows obtaining more robust sampling results. Additionally, a Bayesian prediction framework is derived that allows obtaining a full predictive distribution from which point predictions as well as all other moments of the prediction can be derived. Further, by relying on the Bayesian approach developed in STORM et al. (2015) asynchronous data can be considered directly without the need of interpolating macro data as in previous studies.

The motivation for the paper was a particular need of the European Commission to predict farm numbers in different classes (see the introduction section). In this paper we addressed this need by developing a prediction approach that makes use of two data sources (FSS and FADN) available for the EU. The Bayesian Markov prediction framework consistently combines the two data sources and specifically exploits the advantages of each.

The results of the out-of-sample predictions show that even with the combined information from the two data sources it is very difficult to outperform the constant or linear prediction using a Markov approach. Overall, the three prediction methods perform very similar with no method clearly outperforming one another.

Several conclusions and implications follow from a policy and scientific point of view. First, when applying sophisticated methods to predict farm structural change, their performance in out-of-sample predictions relative to a constant or linear prediction may show limited gains of prediction accuracy in the short term. Secondly, even the additional use of FADN sample data with the developed Bayesian Markov approach may not outperform simpler prediction approaches when the change during the period is rather small or follows a clear trend. Thirdly, with respect to the Markov estimation approach we conclude that a robust estimation approach to consistently combine FSS and FADN data is now available and may improve prediction in other settings where included determinants of non-stationary transition probabilities play a more significant role during the prediction period.

A limitation of the study and the performed prediction is the limited set of explanatory variables included in the Markov approach rendering it likely that important drivers of farm structural changes are missed or measured incorrectly. A further limitation is the broad classification of farms in three size classes and three farm types. Due to the diversity of farms within each class, it is likely that explanatory variable have different effects for individual farm within a class, which further complicates the selection and interpretation of explanatory variables. The number of size classes is predefined by the FADN data and cannot be extended for the time period used in estimation. The number of farm types could be extended in principle but is limited by the increasing number of zero observations for particularly combination of farm type, region, year and size class that arise when further differentiating farm types.

Even though the focus of the paper is a short-term prediction of farm numbers, we like to point out that the employed Markov approach is useful for other purposes as well. The increased data information compared to previous estimation procedures for Markov transition probabilities likely improves the ability to identify marginal effects of determinants, a challenge different from the prediction accuracy considered in this paper.

Literature

BREUSTEDT, G. and T. GLAUBEN (2007): Driving Forces behind Exiting from Farming in Western Europe. In: Journal of Agricultural Economics 58 (1): 115-127.

EPEC (European Parliament and European Council) (2013): Regulation (EU) No 1307/2013 of the European Parliament and of the Council of 17 December 2013 establishing rules for direct payments to farmers under support schemes within the framework of the common agricultural policy and repealing Council Regulation (EC) No. 637/2008 and Council Regulation (EC) No 73/2009.

GOCHT, A., T. HECKELEI, S. NEUENFELDT, N. RÖDER and H. STORM (2013): Modelling the Effects of the CAP on Farm Structural Change. Espinosa, M. and S. Gomez y Paloma (eds.): JRC scientific and technical reports. Publications Office of the European Union, Luxembourg. URL: http://ftp.jrc.es/EURdoc/JRC75524.pdf.

HYNDMAN, R.J. and A.B. KOEHLER (2006): Another look at measures of forecast accuracy. In: International Journal of Forecasting 22 (4): 679-688.

LIU, J.S. (2008): Monte Carlo strategies in scientific computing. Springer, New York.

MORO, D. and P. SCKOKAI (2013): The impact of decoupled payments in farm choices: Conceptual and methodological challenges. In: Food Policy 41: 28-38.

PIET, L., L. LATRUFFE, C. LE MOUEL and Y. DESJEUX (2012): How do agricultural policies influence farm size inequality? The example of France. In: European Review of Agricultural Economics 39 (1): 5-28.

STOKES, J.R. (2006): Entry, Exit, and Structural Change in Pennsylvania's Dairy Sector. In: Agricultural and Resource Economics Review 35 (2): 357-373.

STORM, H., T. HECKELEI and R.C. MITTELHAMMER (2015): Bayesian estimation of non-stationary Markov models combining micro and macro data. In: European Review of Agricultural Economics. Advanced Access, DOI:10.1093/erae/jbv018.

ZIMMERMANN, A. and T. HECKELEI (2012a): Differences of farm structural change across European regions. Discussion Paper 2012:4. Institute for Food and Resource Economics, University of Bonn.

– (2012b): Structural Change of European Dairy Farms – A Cross-Regional Analysis. In: Journal of Agricultural Economics 63 (3): 576-603.

ZIMMERMANN, A., T. HECKELEI and I.P. DOMÍNGUEZ (2009): Modelling farm structural change for integrated ex-ante assessment: review of methods and determinants. In: Environmental Science & Policy 12 (5): 601-618.

Acknowledgements

The work conducted in this paper is in parts funded by the German Research Foundation (DFG), grant No. HE 2854/4-1 as well as the research project "*Modelling the effects of the CAP on farm structural change*" (Contract 151949-2010-A08-DE) from the European Commission Joint Research Centre – Institute for Prospective Technological Studies (IPTS). The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission. The authors are grateful to Norbert Röder, Alexander Gocht and Sebastian Neuenfeldt for helpful comments and support regarding data access and handling.

Contact author:
DR. HUGO STORM
University of Bonn
Institute for Food and Resource Economics
Nußallee 21, 53115 Bonn, Germany
e-mail: hugo.storm@ilr.uni-bonn.de