seamless

System for Environmental and Agricultural Modelling;
Linking European Science and Society

Report and Code to Simulate Structural Change

Zimmermann, A., Heckelei, T., Adenäuer, M.

Partner involved: UBONN

Report no.: 31
November 2007
Ref: PD3.6.10
ISBN no.: 90-8585-119-X
and 78-90-8585-119-6

universitätbonn

Logo's main partners involved in this publication

Sixth Framework Programme

SEAMLESS integrated project aims at developing an integrated framework that allows ex-ante assessment of agricultural and environmental policies and technological innovations. The framework will have multi-scale capabilities ranging from field and farm to the EU25 and globe; it will be generic, modular and open and using state-of-the art software. The project is carried out by a consortium of 30 partners, led by Wageningen University (NL).

Email: seamless.office@wur.nl
Internet: www.seamless-ip.org

Authors of this report and contact details

Name: Andrea Zimmermann — Partner acronym: UBONN
Address: Nussallee 21, 52115 Bonn
E-mail: andrea.zimmermann@ilr.uni-bonn.de

Name: Thomas Heckelei — Partner acronym: UBONN
Address: Nussallee 21, 52115 Bonn
E-mail: Thomas.heckelei@ilr.uni-bonn.de

Name: Marcel Adenäuer — Partner acronym: UBONN
Address: Nussallee 21, 52115 Bonn
E-mail: marcel.adenaeuer@ilr.uni-bonn.de

When citing this SEAMLESS report, please do so as:

Zimmermann. A., Heckelei, T., Adenäuer, M., 2007. Report and Code to Simulate Structural Change, SEAMLESS Report No.31, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 49 pp, ISBN no. 90-8585-119-X and 78-90-8585-119-6.

Table of contents

General information

Task(s) and Activity code(s):	Task 3.6, Activity 3.6.5
Input from (Task and Activity codes):	Task 3.6, Activity 3.6.5
Output to (Task and Activity codes):	Task 3.6, Activity 3.6.5
Related milestones:	M3.6.5

Executive summary

The main purpose of this deliverable is to document and explain the methodology applied for the estimation of structural change and the link of the structural change module to SEAMLESS-IF. As the structural change module is still under development this deliverable is intended to document the current state of the model. Future development steps regarding the methodology are however included. Results for the methodologically improved model will be shown in updated versions of the deliverable which are due in month 39 and as a final version in 2008.

For the sake of completeness and to give the reader a better understanding of the process of structural change and difficulties concerning the data and typology the chapters “Structural Change in Agriculture” and “Data and Typology” are taken from the preceding deliverables PD3.6.6 and PD3.6.7 and repeated here in a summarized version. The main part of the deliverable comprises the methodological explanation of the Markov chain approach. As the structural change module is still under development only a part of the approach described could already be examined in terms of results. Hence, preliminary results from a first model version are shown for a number of test regions. The last chapter deals with the question on how to link the structural change module to SEAMLESS-IF.

Specific part

1 Introduction

The structural change module in SEAMLESS-IF is used to retrieve time-adjusted aggregation weights to establish the regional coverage in the up-scaling procedure from the farm to the market level. The aggregation weights are found applying a Markov chain estimation where the past development of farm numbers in certain farm types is used to derive transition probabilities which in turn are used to forecast the future farm numbers in these farm types. Thus, the process of structural change is represented by farm numbers only, i.e. the structural change module is not able to capture effects that take place internally on the farms. These changes are supposed to be represented in the FSSIM models. The impact of structural change on the agricultural sector performance will be examined with CAPRI.

As shown in Zimmermann et al. 2006 the Markov chain technique in agricultural economics has so far been applied for only one region or country without regional differentiation and one agricultural specialisation (e.g. dairy farming) per analysis. In the analysis at hand the attempt is made to cover the whole agricultural sector through differentiation between various specialisation classes and, as a second innovation, a cross-regional analysis is conducted in order to take into account different development patterns throughout the European Union.

For the sake of completeness and to give the reader a better understanding of the process of structural change and difficulties concerning the data and typology the chapters “Structural Change in Agriculture” and “Data and Typology” are taken from the preceding deliverables PD3.6.6 and PD3.6.7 and repeated here in a summarized version. The main part of the deliverable comprises the methodological explanation of the Markov chain approach. As the structural change module is still under development only a part of the approach described could already be examined in terms of results. Hence preliminary results from a first model version are shown for a number of test regions. The last chapter deals with the question on how to link the structural change module to SEAMLESS-IF.

2 Structural change in Agriculture

To start with, a few theoretical considerations regarding the process of structural change in agriculture are made which refer to the definition and the potential driving factors of structural change. Especially the latter become important when it comes to choosing explanatory variables for the modelling part of the analysis.

2.1 Definition of structural change

One of the problems faced in the analysis of structural change in agriculture is the heterogeneity of the definition of *farm structure*. There is basically a general recognition of the complexity of this concept but no single widely accepted definition (Stanton 1993, Balmann 1997). Balmann, for example, defines it as: "who is producing what, in what amounts and by what means?" (Balmann 1997, p. 106). Nevertheless, from a wider perspective the concept of agricultural structure can be framed by looking at its main elements: farm size, resource ownership and control, managerial and technological requirements, tenure pattern, importance of part-time operations, degree of vertical integration in a given industry, organisation of production, ease of entry into farming as an occupation and manner of asset transfer to succeeding generations (Penn 1979, Tweeten 1984, Knutson et al. 1990). This list is not exhaustive but pretends to cover the main definitory elements of agricultural structure found in the literature.

The definition of *structural change* varies depending on the underlying definition of the *agricultural structure*. Basically there are two orientations: one relating to productivity changes (e.g. Oehmke et al. 2004, Kim et al. 2005) and another relating to the structure of the industry. The first definition of structural change leads to the wide field of time series analyses (e.g. determination of structural breaks) which is extensively covered in the branch of general economics. In agricultural economics, however, the focus of the discussion often lies on changes in the structure of the industry. Nevertheless, in most studies both are evaluated together, since farm structure is usually not independent of production relationships. In the SEAMLESS context structural change can simply be defined as the change of the number of farms in different farm types (as classified e.g. according to different size or activity measures, age cohorts, specialisation classes etc.).

2.2 Factors contributing to structural change

Most studies on farm structures provide an enumeration of the factors assumed to determine structural change in agriculture. Here, only a short overview of these factors is given, leaving the in-depth discussion to others (e.g. Hallam 1991, Hallam 1993, Goddard et al. 1993, Harrington et al. 1995).

Factors which are broadly assumed to affect the process of structural change are the ones listed and briefly explained below. These factors should not be seen as mutually exclusive but rather interrelated with each other (U.S. Congress 1985, Goddard et al. 1993, Harrington et al. 1995, Hallam 1991, Boehlje 1990).

1) The *technology* model is based upon the concepts of economies of scale and the adaptation and diffusion of technology. The literature on economies of size has focused fundamentally on the long run cost curve in agricultural production and the determinants that shape and shift that curve (Boehlje 1990). The adaptation and diffusion of technology refers to the concept of Cochrane's treadmill (Cochrane 1958). The con-

cept focuses on the impact of technological innovation reducing real per unit cost of output at the farm level and with competition encouraging farmers to adopt new technologies. The first adopters of the new technology will gain from the first-mover advantage (Bremmer et al. 2004), but as adoption becomes widespread, prices of farm commodities will fall differently per farm size, triggering structural adjustments (Ahearn et al. 2002).

2) *Off-farm employment* is handled in two ways. On the one hand, it could be seen as a first step out of the sector. As opportunity costs increase due to better wage levels outside of agriculture, farmers tend to leave the sector until wages equalize (Hallam 1991) or try to achieve comparable incomes by enlarging the farm business (Harrington et al. 1995). On the other hand, off-farm employment provides a method to keep on farming at small scales if the off-farm income complements the household income (Goddard et al. 1993) or farmers are even willing to subsidize their small farm at least in the short-run from other income sources (Harrington et al. 1995).

3) *Public programs* are governmental policies that impact the agricultural sector in different ways according to their design. Examples often mentioned are tax policy, commodity programs, credit programs, general monetary and fiscal policies, and public research and extension efforts (Harrington et al. 1995, Goddard et al. 1993, U.S. Congress 1985).

4) *Human capital* refers to and is influenced by the managerial capability, level of schooling, public education programs. It is assumed that an increase in human capital allows the firm manager to more effectively process information used to allocate the firm's resources and to evaluate new technologies. Thus, an increase in human capital allows for effectively managing an increasing firm size (Goddard et al. 1993).

5) *Demographics* refer mainly to the age structure of farm operators and the shrinking number of entrants to the farming sector. One might argue that these aspects are a consequence rather than a cause of structural change. However, the speed of change in a region might be heavily influenced by the age structure of the farmers. Goddard et al. (1993) also point to the changes in the demographical structure of the general population that might have some influence concerning the demand of agricultural products.

6) *Market structure* itself influences structural change. This point is derived from the industrial organization structure (Boehlje 1990). The way in which prices are set is determined by the nature of the market, so that the conduct of the industry is a function of its structure (polypoly, oligopoly, monopoly vs. polypsony, oligopsony, monopsony). The development of institutional arrangements, such as vertical integration and cooperatives has an (so far unclear) impact on structural change as well (Goddard et al. 1993).

7) *Economic forces.* Several sector specific and macroeconomic factors such as input and output prices, demand changes, and the interest rate are supposed to have an impact on structural change (Hallam 1991, Goddard et al. 1993). However, most of the aforementioned points could be expressed in economic terms as well such that economic factors could in fact be seen as the heading under which the other factors are summarized.

3 Data and Typology

This chapter describes the SEAMLESS typology and its use in the structural change module. Thereafter data sources and limitations are discussed.

3.1 SEAMLESS typology

The SEAMLESS typology consists of four dimensions which are used for a mutually exclusive grouping of farms into farm types. These are the size dimension (3 classes), the specialisation dimension (10 classes), the land-use dimension (9 classes) and the intensity dimension (3 classes). By combination of the four dimensions and taking out unneeded farm types one arrives at a total amount of 189 farm types. As this amount is technically not manageable in the Markov chain estimation (the maximum number of farm types that was dealt with in the literature was 18 farm types, compare Zimmermann et al. 2006) only two of the four dimensions are used in the structural change module. These are the size and the specialisation dimension.

The size dimension as defined by SEAMLESS contains three economic size categories: a small category until 16 ESU, a medium category from 16 to 40 ESU and a large category greater or equal to 40 ESU (Table 1). Economic size categories are chosen because they allow the comparison of farms across different specialisation classes.

Table 1: Types in the size dimension and definitions

Size type	Definition
Small scale	< 16 European size units (ESU)
Medium scale	≥ 16 ESU and < 40
Large scale	≥ 40 ESU

Source: Andersen et al. 2006.

The specialisation dimension comprises 10 categories which correspond to the official Community 'Types of farming' as shown in Table 2. The codes used in the structural change module are shown in parentheses in the first column. The size classes are then indicated by an underscore followed by the symbol 'S' for small, 'M' for medium and 'L' for large farms (e.g. ARAB_S means small arable farms).

Table 2: Types in the specialisation dimension with definitions and reference to codes in the Community typology

Specialisation type	EU-Code	Definition
Arable systems (ARAB)	1 + 6	>2/3 of SGM from arable or (>1/3 of SGM from arable and/or permanent crops and/or horticulture)
Dairy cattle (DARY)	4.1	>2/3 of SGM from dairy cattle
Beef and mixed cattle (BEEF)	4.2 + 4.3	>2/3 of SGM from cattle and <2/3 of SGM from dairy cattle

Sheep, goats and mixed grazing livestock (SHGM)	4.4	>2/3 of SGM from grazing livestock and <2/3 of SGM from cattle
Pigs (PIGS)	5.01	>2/3 of SGM from pigs
Poultry and mixed pigs/poultry (POLT)	5.02 + 5.03	>2/3 of SGM from pigs and poultry and <2/3 of SGM from pigs
Mixed farms (MIXF)	8	All other farms
Mixed livestock (MIXL)	7	>1/3 and <2/3 of SGM from pigs and poultry and/or >1/3 and <2/3 of SGM from cattle
Permanent crops (PERM)	3	>2/3 of SGM from permanent crops
Horticulture (HORT)	2	>2/3 of SGM form horticultural crops

Source: Andersen et al. 2006.

The combination of these two dimensions results in a total amount of thirty farm types.

3.2 Data

Two EU databases contain information on the number of farms in different farm classes that is necessary for the structural change module: the Farm Structure Survey (FSS) and the Farm Accountancy Data Network (FADN).

The FSS database contains information on the structure of agricultural holdings collected through agricultural structure surveys. The variables are arranged into four groups: a general one with the key variables, and three specialized ones containing detailed data on land use, livestock, management and farm labour input. One distinguishes between basic and intermediate surveys. The basic surveys are in line with the Food and Agriculture Organization (FAO) recommendations on a world-wide agricultural census and are carried out every 10th year. The intermediate surveys are organised three times between censuses. All surveys relate to crop years and the exact reference periods are determined in special legislations. Whereas the basic surveys are generally full scope censuses, the intermediate surveys are conducted on a random sample base. The sampling rate of the latter depends on the country and the year of survey. It varies between 3-40% of the total population of agricultural holdings. In some Member States every survey is census. FSS data are available for the years 1990 (1989/90), 1993, 1995, 1997, 2000 (1999/2000), and 2003 (European Commission 2000, European Commission 2006).

FADN is based on sample farms only. Unlike the FSS which relates to NUTS2 regions, FADN farms are spatially grouped according to special FADN regions which are mostly aggregates of the NUTS2 regions. However, in a limited number of cases the borders of FADN and NUTS2 do not match exactly. At EU-25 level NUTS2 comprises 281 European regions, whereas the FADN database distinguishes only 120 regions. FADN depends on the FSS in the sense that the FSS is used to derive aggregation weights for extrapolating FADN sample data to the whole EU. The FADN database contains not only aggregated but detailed accountancy data for each sample farm. Each sample farm represents a number of similar farms in the same region. A special weighting system is used in the calculation of the FADN results. It is based on the principle of "free expansion": for each holding in the sample, an individual weight is applied (extrapolating factor). In order to calculate this individual weight, holdings in the sample and in the field of survey are stratified according to the same three criteria: FADN region, type of farming and economic size class. The individual weight is equal to the ratio between the numbers of holdings of the same classification cell (FADN region x type of farming x economic size class) in the population (FSS data) and in the sample (FADN data). Regarding the selection of FADN sample farms it is current practice for the national Liaison Agencies to design their own selection plans for the European Union survey. The plans are submitted to the FADN European Union Committee for approval and can vary in technical sophistication from one Member State to another (European Commission 2005). FADN sample farms may be surveyed several years in succession, but are usually replaced by a similar farm after 5-7 years such that time series are not available for a specific set of farms. FADN data is provided annually. As the FSS is not annual but every 2-3 years, the Commission services use the most recent information available for extrapolating the FADN sample data. Which FSS years are representative for the years without survey is shown in Table 3.

Table 3: FADN-FSS sampling years

FSS	1990		1993		1995		1997			2000			2003	
FADN	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003

Source: European Commission 2005.

The total number of farms in the different stratification criteria which should be met by the number of farms represented by FADN (the extrapolated FADN sample farms) thus does not vary annually but according to FSS updates. Another difference between FADN and FSS is that in FADN only ‘professional’ farms are represented whereas the FSS includes ‘non-professional’ farms as well. ‘Professional’ farms are defined as farms whose economic size exceeds a certain threshold. The size thresholds vary from Member State to Member State and have been adjusted over time as well. They are given in the appendix (Table 12). The occurring differences regarding the coverage of certain characteristics of European agriculture between the FADN representative data and FSS are shown in Table 4.

Table 4: Differences FADN-FSS (year 2000)

Member State	Number of farms		Coverage field of observation FADN			
	FSS	FADN	Farms [%]	ESU [%]	UAA [%]	AWU [%]
Belgium	61710	39350	63.8	96.2	92.1	84.2
Denmark	57830	45610	78.9	98.3	96.2	94.1
Germany	471960	296740	62.9	96.9	93.9	87.0
Greece	817060	531640	65.1	94.6	92.6	89.8
Spain	1287420	873030	67.8	97.2	88.2	86.8
France	663810	437850	66.0	97.9	95.6	89.5
Ireland	141530	123190	87.0	99.4	96.2	92.9
Italy	2153720	1127200	52.3	95.4	91.5	77.9
Luxembourg	2810	1970	70.1	95.9	95.2	87.2
The Netherlands	101550	79160	78.0	97.8	94.4	91.8
Austria	199470	81880	41.0	86.3	62.0	62.9
Portugal	415970	312795	75.2	96.0	95.4	82.8
Finland	81190	49790	61.3	94.0	84.2	82.6
Sweden	81410	41010	50.4	93.9	86.1	80.1
United Kingdom	233250	124610	53.4	98.0	88.0	80.0

Source: European Commission 2005.

As can be seen from the table the total number of farms varies considerably between the two databases. The FADN database covers only a range of about 40% (Austria) to 87% (Ireland) of the farms that are represented in the FSS but achieves a coverage of the total economic size of the national agricultural sectors of about 86.3% (Austria) to 99.4% (Ireland). The coverage of the total UAA of the FADN sample farms lies between 62.0% (Austria) and 96.2% (Denmark and Ireland). In terms of total Average Working Units (AWU) a coverage of 62.9% (Austria) and 94.1% (Denmark) is reached.

Regarding the data needs for the structural change module it was initially thought of combining the FSS and FADN data somehow. However, the FSS data turned out to be not available in the detail necessary to meet the SEAMLESS typology demands as a result of which the structural change module is based on FADN data only now.

4 Estimation procedure

The chapter is divided into three parts. Firstly, the basic concept of the Markov chains is explained, followed by the description of the estimation procedures for stationary and non-stationary transition probabilities, respectively.

4.1 Markov chains

The estimation of Markov chains has a long tradition in the analysis of structural change in agriculture and is a widely accepted approach to predict the number of farms in certain farm types (Zimmermann et al. 2006, Zepeda 1995a, Karantininis 2002, Stavins et al. 1980).

In a Markov chain the movement of firms from a specific firm category (e.g. a farm type) to another one is seen as a stochastic process which can be represented by transition probabilities. Usually, the movement of farms between several farm types is supposed to follow a first order Markov chain, i.e. it is assumed that the probability of the movement of a farm at time t to another farm type in the period $t+1$ is independent of earlier periods:

$$P\{s_t = j | s_{t-1} = i, s_{t-2} = k, ...\} = P\{s_t = j | s_{t-1} = i\} = p_{ij},$$

where s_t $\{1, 2, ..., S\}$ is a discrete, stochastic variable and i and j are the states (farm types) a specific farm can be in. The transition probability p_{ij} represents the probability of a movement from state i to state j. The single transition probabilities can be summarized in a transition probability matrix P $(S \times S)$:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1S} \\ p_{21} & p_{22} & \cdots & p_{2S} \\ \vdots & \vdots & \cdots & \vdots \\ p_{S1} & p_{S2} & \cdots & p_{SS} \end{bmatrix}.$$

If micro-data is available, i.e. data from which the exact number of movements from one farm type to another can be derived, the equation to be estimated becomes:

$$p_{ij} = \frac{m_{ij}}{\sum_{j=1}^{S} m_{ij}}, \qquad (0.1)$$

where m_{ij} denotes the number of movements of firms from state i to state j during the time period under discussion and S is the total number of states. have shown that the above given approximation of the true p_{ij} is, in fact, the maximum likelihood estimate. However, in most cases as detailed farm data as required for the micro-data estimation approach is not available and one has to rely on more aggregated data (macro-data). The Markov chain can then be estimated using the shares of farms in the different farm types:

$$y_{j(t)} = \sum_{j=1}^{S} y_{i(t-1)} p_{ij}, \qquad (0.2)$$

where $y_{j(t)}$ denotes the share of farms in farm type j at time t. In the Markov chain literature $y_{(t-1)}$ is equally often described as x_t in order to avoid the cumbersome subscripts and come to the more familiar standard regression equation format. Since probabilities are not allowed to be less than zero and the process has to result in some state it follows that:

$$\sum_{j=1}^{S} p_{ij} = 1 \quad (0.3)$$

and

$$p_{ij} \geq 0 \quad (0.4)$$

for $(i = 1, 2, ..., S)$.

From the transition probabilities predictions on future farm numbers in any state can be easily calculated:

$$X_t = X_0 P^t, \quad (0.5)$$

where X_0 is the initial starting state vector or the initial configuration of individuals in the S states, where x_{0i} represents the number of individuals in state i during time period $t = 0$, and X_t is the t^{th} configuration vector.

One of the strongest assumptions of the Markov model is that the transition probabilities do not change in the whole process, i.e. they are said to be stationary. This implies that the process of structural change follows the same path until an equilibrium solution is achieved. This may represent a realistic assumption as long as all other factors remain the same, too. However, this assumption does not hold for most economic phenomena. Changes in exogenous variables, e.g. wages, prices, technology or policy, require the determination of non-stationary (time-varying) transition probabilities with an econometric model "behind" the pure Markov chain. The non-stationary transition probabilities are, hence, specified as functions of exogenous variables and parameters:

$$p_{ij(t)} = f_{ij}(Z_{(t-1)}, \beta_{ij}) \quad (0.6)$$

where f_{ij} is the function of the vector of (lagged) explanatory variables Z and the vector of parameters β_{ij} which relates the exogenous variables to the transition probabilities. The explanatory variables appear time-lagged because they are related to transitions which begin at time $t-1$.

4.2 Stationary transition probabilities

In our case does the TPM consist of 31 x 31 transition probabilities (resulting from the 30 farm types described above plus the artificial entry/exit class), that means 961 parameters are to be estimated. The data available comprises 6 observation years which are unevenly spread over the time period 1990-2003 (1990, 1993, 1995, 1997, 2000 and 2003). Hence, only 155 observations (5 transitions x 31 farm types) are available for the estimation of 961 parameters. There are basically three alternatives on how to deal with the severe lack of degrees of freedoms resulting from the limited amount of data available.

1. Interpolation of data points between the FSS years.
2. Diminishment of the problem.
3. Application of non-traditional estimation techniques which allow for the incorporation of *a priori* information.

In case of the stationary transition probabilities a combined approach has been chosen where the data is interpolated and the problem is reduced to the estimation of transition probabilities for transitions between size classes only, as a result of which a rather simple least-squares estimation technique could be applied. The data interpolation has the additional advantage of a facilitated interpretation of the transition probabilities as these can now be understood as annual probabilities. In a further step the extension of the problem through the estimation of probabilities for transitions between specialisation classes as well is planned. This will be carried out by applying a cross-entropy approach. Both approaches, least-squares and cross-entropy, are described below.

4.2.1 Least-squares

Least-squares estimates of p_{ij} are obtained by adding an error term e_{jt} to equation (0.2) and choosing a set of estimates $\hat{p}_{ij}$ so to minimize $\sum_t \hat{e}_{jt}^2$ subject to the probability constraints (summing up to unity (0.3) and non-negativity (0.4)):

$$\hat{e}_{jt} = y_{j(t)} - \sum_{i=1}^{S} y_{i(t-1)} \hat{p}_{ij} \tag{0.7}$$

However, from an econometrical point of view one has to remark critically that the least-squares estimates of the macro-data model are in fact inconsistent since y is actually multinomially distributed (MacRae 1977). The least-squares technique is nonetheless applied here because it turned out to be the only procedure which could solve the problem without making rather strict assumptions regarding the underlying processes or the help of *a priori* information and important insights helping to understand the process of structural change could be gained anyway.[1]

To receive a model that is tractable with the least-squares approach the amount of parameters has to be reduced. This is done by restricting the problem to the estimation of transition probabilities between size classes within each specialisation class only. The probabilities for changes between specialisation classes are assumed to be zero. Also, probabilities for movements into farm types where during the whole observation period no farms were represented are set to zero (e.g. there are no farms in the small size category in the Netherlands). The single probabilities for each specialisation are connected via the entry/exit class and thus not totally independent from each other. That means that although no probabilities for direct movements between specialisations are estimated, increasing shares in one and decreasing shares in another specialisation class are accounted for through increasing probabilities for exit in the first and increasing probabilities for entry in the latter specialisation. See Table 10 for an example.

[1] The theoretically most valid approach would have been the estimation of a multinomial logit model (Zepeda 1995b, Zepeda 1995a, MacRae 1977). However, this was not possible because the number of farm types was too large. The multinomial logit has been applied in the literature up to an amount of four farm types (i.e. 16 transition probabilities).

This is the simplest and most straightforward version of the Markov chain model. Since changes between specialisation classes are not allowed degrees of freedom could be saved and only 151 transition probabilities had to be estimated for every region[2].

4.2.2 Cross-entropy

A suitable approach for the estimation of such ill-posed problems as given here is the incorporation of *a priori* information. Lee et al. (1996) and Golan et al. (1996) developed a maximum entropy procedure for the estimation of stationary Markov processes. Here a Generalized Cross Entropy (GCE) formulation according to Lee et al. (1996), Golan et al. (1996) and Karantininis (2002) is presented. The GCE formalism minimizes the distance between the probabilities that are consistent with the data and the prior information and as such differs from the case of restricted estimators where constraints must always hold. Here, the prior information can be overruled by the information coming from the sample data.

The GCE Markov problem can be stated as follows:

$$\min H(p_{ij}, q_{ij}, w_{itm}, u_{itm}) = \sum_{i}^{K}\sum_{j}^{K} p_{ij}\ln(p_{ij}/q_{ij}) + \sum_{i}^{K}\sum_{t}^{T}\sum_{m}^{M} w_{itm}\ln(w_{itm}/u_{itm}) \quad (0.8)$$

subject to the Markov consistency constraint:

$$\begin{aligned} &\mathbf{y}_T = (I_K \otimes X_T)\mathbf{p} + \mathbf{e}_T, \\ &(TK \times 1) = (TK \times K^2)(K^2 \times 1) + (TK \times 1) \end{aligned} \quad (0.9)$$

with

$$e_{it} = \sum_{m}^{M} v_m w_{itm} \quad (0.10)$$

and

$$\sum_{j}^{K} p_{ij} = 1 \quad (0.11)$$

$$\sum_{m}^{M} w_{itm} = 1 \quad (0.12)$$

$$p_{ij} \geq 0 \quad (0.13)$$

$$w_{itm} \geq 0 \quad (0.14)$$

Equation (0.8) represents the GCE function which minimizes the distance between the data in the form of Markov transition probabilities p_{ij} and the Markov transition priors q_{ij}. By analogy, the GCE algorithm minimizes also the distance between the error in the form of posterior probabilities w_{itm} and the priors u_{itm}. Equation (0.9) represents the Markov data consis-

[2] The 151 transition probabilities to be estimated arrive from the 10 specialisation classes times 9 transition probabilities in each class resulting from the 3 size classes plus 61 probabilities for entry respectively exit.

tency constraint or moment condition, where $\mathbf{y}_T$ are the elements of a $TK \times 1$ vector of known proportions falling in the k-th Markov state in time $t+1$, $(I_K \otimes X_T)$ is the Kronecker product of a $K \times K$ identity matrix and the matrix of proportions in the K states in time t X_T. The Markov transition probabilities p_{ij} directly enter the GCE objective function (0.8) without needing to be parameterized.[3] The error term $\mathbf{e}_T$ is parameterized as given by equation (0.10) following the Shannon's entropy formulation, where v is an M-dimensional vector of support points and w is an M-dimensional vector of weights (in the form of probabilities) for each e_{it}. Several authors suggest setting the support vector to $\mathbf{v} = [-1/K\sqrt{T},...,0,...,1/K\sqrt{T}]'$ (Karantininis 2002, Tonini 2007). The equations (0.11) and (0.12) represent the adding-up constraints for the transition probabilities and error weights, respectively. Equations (0.13) and (0.14) impose the non-negativity constraints on the probabilities.

Generally, concerns are expressed regarding the adding of subjective information to the model through the somewhat intransparent use of *a priori* information in GCE approaches. This applies for the Markov chain estimation as well, even though at lower rates because not supports and weights but point priors are set for the transition probabilities. In this sense the GCE Markov model might even be superior to traditional estimation techniques because here the prior information is allowed to be overruled if not supported by the data whereas in most traditional Markov chain applications the incorporation of strict assumptions was necessary in order to come to meaningful estimates. However, in order to meet the concerns regarding the incorporation of subjective opinions ('expert knowledge') it is envisaged to use information available from the single farm data as *a priori* information for the macro data model in further development steps of the Markov chain approach. A straightforward way to do this would be the estimation of a micro data Markov model according to equation (0.1) and then use the resulting maximum likelihood estimates as priors for the transition probabilities ((Golan et al. 1996), p. 59).[4] However, this appears to be difficult in our case because the individual FADN sample farms change arbitrarily and there is no information on exits available. To bring together information from both the macro and the micro data a sampling procedure could be used before estimating the Markov model, but this remains to be worked out in future versions of this deliverable.

4.3 Non-stationary transition probabilities

There exist two fundamentally different approaches to deal with non-stationary transition probabilities in the literature: with traditional estimation techniques 'real' non-stationary transition probabilities can be estimated, i.e. coefficients for the effect of each explanatory variable on the transition probabilities are estimated and a different TPM is retrieved for every year (equation (0.6)). In GCE approaches only instrumental variable techniques have been applied so far where the effect of explanatory variables can be recovered through the calculation of transition probability elasticities but no coefficients are estimated and the resulting

[3] Here, Karantininis (2002) differs from the approach of Lee et al. (1996) and Golan et al. (1996) who suggest replacing the point prior with a prior that permits a discrete probability distribution to be specified for each of the p_{ij} so that $p_{ij} = \sum_M z_m p_{mij}$, with z_m being a vector of supports.

[4] Alternatively, information from the macro model could be used to estimate the micro model.

TPM does not vary over time[5]. However, theoretically it would be possible to estimate a 'real' non-stationary model with GCE as well. This non-stationary approach is described in chapter 4.3.1, the instrumental variables approach is described in chapter 4.3.2.

4.3.1 Non-stationarity

In case of the non-stationary Markov chain approach the p_{ij} are assumed to follow the relationship:

$$p_{ij}(t) = f_{ij}(\mathbf{z}_{ij}(t), \beta_{ij}) + e_{ij}(t), \tag{0.15}$$

where $f_{ij}(\cdot)$ is a function relating each element $p_{ij}(t)$ of the non-stationary transition probability matrix (NSTPM) to a vector of explanatory variables $z_{ij}(t)$. The β_{ij} are parameters of the $f_{ij}(\cdot)$, and $e_{ij}(t)$ is the disturbance term. The Markov process can now be expressed as:

$$\mathbf{y}(t+1) = \mathbf{x}'(t)[\boldsymbol{\beta}\mathbf{z}(t) + \mathbf{e}(t)] + \mathbf{u}(t). \tag{0.16}$$

Each β_{ijn} and each e_{ijt} can be parameterised over a discrete finite support space: $\beta_{ijn} = \sum_s^S d_{ijns}\theta_s$, and $e_{ijt} = \sum_h^H g_{ijth}\varphi_h$, where $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$ are support vectors of size S and H, respectively, and $\mathbf{d}$ and $\mathbf{g}$ are the corresponding probabilities to be recovered. The Markov process in (0.16) now becomes:

$$y_{jt} = \sum_i^K x_{it}\left[\sum_n^{N_{ij}}\left(\sum_s^S d_{ijns}\theta_s\right)z_{nt} + \sum_h^H g_{ijth}\varphi_h\right] + \sum_m^M v_m w_{jtm}, \tag{0.17}$$

where N_{ij} is the number of covariates in the (ij)th cell. Applying GCE the $\boldsymbol{\beta}$, $\mathbf{e}$, and $\mathbf{u}$ can be recovered through the recovered values of $\mathbf{d}$, $\mathbf{g}$, and $\mathbf{w}$, respectively. In order to meet the probability conditions additional constraints need to be imposed on the $\mathbf{d}$. Alternatively, a multinomial Logit transformation could be assumed, which satisfies both the normalisation and the non-negativity constraints automatically but was already said to be inapplicable in the case of more than four Markov states (MacRae 1977, Zepeda 1995b, Zepeda 1995a).

4.3.2 Instrumental variables generalised cross-entropy estimator

The GCE estimator for the NSTPM shown here is similar to the estimator used to derive the social accounting matrix in Golan et al. (2000) and was first applied in the context of a Markov chain estimation by Karantininis (2002).

The information on the covariates Z_{tn}, a $T \times N$ matrix of N covariates in the T time periods, can be incorporated into the GCE model by multiplying both sides of the data consistency constraint (0.9) with Z_{tn}:

[5] Although the transition probabilities do not vary over time they are called 'non-stationary' meaning that they are estimated depending on other explanatory variables.

$$\sum_{t}^{T} z_{tn} y_{tj} = \sum_{t}^{T} \sum_{i}^{K} z_{tn} x_{it} p_{ij} + \sum_{t}^{T} \sum_{m}^{M} z_{tn} v_m w_{jtm} \; . \qquad \forall j = 1,\dots,K; n = 1,\dots,N \tag{0.18}$$

Priors are introduced in the objective function (0.8) in the form of matrices $\mathbf{Q}$ (corresponding to the transition probabilities $\mathbf{P}$) and $\mathbf{U}$ (for the disturbance probabilities $\mathbf{W}$). The priors $\mathbf{U}$ are assumed uniformly distributed around zero, hence they add no additional information to the model. The solution to this problem is:

$$\tilde{p}_{ij} = \frac{q_{ij} \exp\left[\sum_t \sum_n x_{it} z_{tn} \tilde{\lambda}_{nj}\right]}{\sum_j q_{ij} \exp\left[\sum_t \sum_n x_{it} z_{tn} \tilde{\lambda}_{nj}\right]}, \tag{0.19}$$

where $\tilde{p}_{ij}$ and $\tilde{\lambda}_{nj}$ are the recovered probabilities and Lagrange multipliers, respectively. Applying the procedure (0.18) not time-varying, but constant transition probabilities are obtained. The effect of the covariates on the transition probabilities can be recovered through the calculation of so-called ‘transition probabilities’ (Zepeda 1995b, Zepeda 1995a). The effect of each z_{tn} on the p_{ij} can be calculated by applying the formula (Karantininis 2002, Tonini 2007):

$$E^{P}_{ijtn} = \frac{\partial \tilde{p}_{ij}}{\partial z_{tn}} \frac{\overline{z}_n}{\tilde{p}_{ij}} = \overline{x}_i \overline{z}_n \left[\tilde{\lambda}_{nj} - \sum_{j}^{K} \tilde{p}_{ij} \tilde{\lambda}_{nj} \right]. \tag{0.20}$$

Similarly, the following elasticity measures the cumulative effect of a unit change in each exogenous variable on the vector of proportions falling into the *k*th Markov state in time (t+1):

$$E^{y(t+1)}_{jn} = \frac{\partial y_{j(t+1)}}{\partial z_{tn}} \frac{\overline{z}_n}{\overline{y}_j} = \frac{\overline{z}_n}{\overline{y}_j} \sum_{i}^{K} \left[\tilde{p}_{ij} \overline{x}_i^2 \left(\tilde{\lambda}_{nj} - \sum_{j}^{K} \tilde{p}_{ij} \tilde{\lambda}_{nj} \right) \right]. \tag{0.21}$$

Applying the constant transition probabilities for the farm number prediction would lead to constant change rates which are not further influenced by exogenous variables and would thus behave similarly to the prediction in case of stationary transition probabilities. However, equation (0.19) allows for updating the transition probabilities once forecasts on the covariates are available.

It is chosen to follow the instrumental variables approach due to three reasons: The methodology is well known and has been applied by several authors now (Karantininis 2002, Jongeneel et al. 2005, Tonini 2007), the use of a priori information is much more transparent than in the non-stationary approach explained in chapter 4.3.1 and it is computationally much simpler and time saving.

5 Preliminary results

Results are shown for the stationary model described in chapter 4.2.1, where probabilities are estimated for the transition between size classes within each specialisation class only and transitions between specialisation classes are left out. Running the simplified Markov model for all regions and all countries takes about one week with the least-squares procedure. In an updated version of this deliverable, results will be shown for a non-stationary model and transitions between specialisation classes, too.

For illustrative purposes results are shown for a number of regions.[6] To give a first picture and before going into detail regarding the regional developments some general results for Germany are given. Germany is one of the countries where remarkable changes in the agricultural sector took place during the last decade and there is an ongoing development with decreasing farm numbers in most regions and across all farm types. It was however expected that especially the small and perhaps medium size classes are obliged to shrinkages whereas the numbers of larger farms would stay relatively stable. As such it came as a surprise when in the transition probability matrices (TPM) for many regions the highest probabilities for exit occurred in the largest size classes, whereas no exits at all are predicted for the small and medium size classes. In the TPM for Northrhine-Westfalia (NRW, Table 10), for instance, this can easily be observed for just the most important specialisation types ‘arable’, ‘dairy’ and ‘mixed farms’. In Table 5 the size classes per specialisation class with the highest exit probabilities are indicated by a cross and compared among all German FADN regions.

[6] Results for the other EU-15 regions will be stored in the SEAMLESS database (see discussion in chapter 6.1) or are available upon request.

Table 5: Size classes with the highest probability to exit, Germany

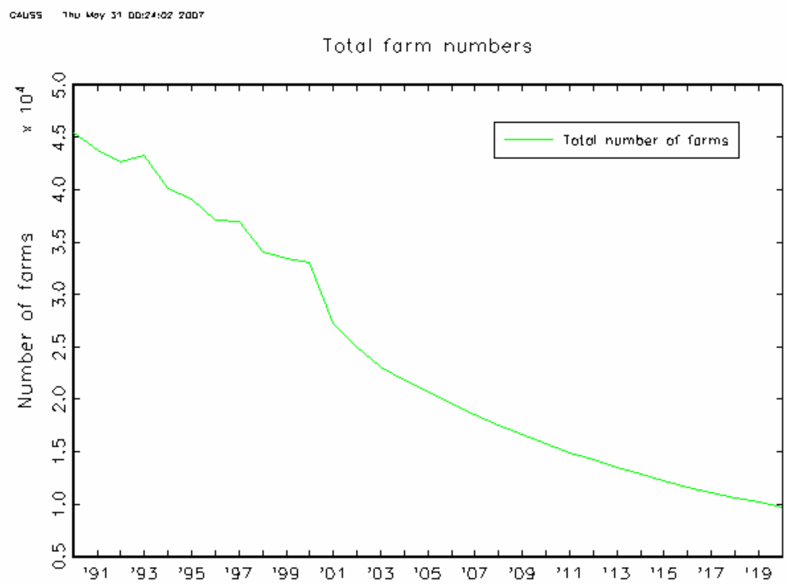
	Schleswig-Holstein	Niedersachsen	Nordrhein-Westfalen	Hessen	Rheinland-Pfalz	Baden-Württemberg	Bayern	Saarland	Brandenburg	Mecklenburg-Vorpommern	Sachsen	Sachsen-Anhalt	Thüringen
ARAB_S									x		x		x
ARAB_M										x		x	
ARAB_L	x	x	x	x	x	x	x	x					
SHGM_S									x				
SHGM_M												x	
SHGM_L													
PERM_S					x								
PERM_M													
PERM_L											x		
DARY_S													
DARY_M													x
DARY_L	x	x	x	x	x	x	x	x			x		
BEEF_S	x	x	x			x			x				
BEEF_M					x			x					
BEEF_L				x							x		
PIGS_S			x	x	x	x							
PIGS_M													
PIGS_L	x												
POLT_S													
POLT_M													
POLT_L													
MIXF_S	x	x						x	x			x	
MIXF_M											x		x
MIXF_L			x	x	x	x	x						
MIXL_S											x		
MIXL_M													
MIXL_L	x	x	x	x			x						
HORT_S		x	x						x		x	x	x
HORT_M													
HORT_L	x											x	

Source: Estimated.

Surprisingly, the table reveals a clear pattern with high probabilities to exit in the large size classes in the West German regions and a more uneven picture for the East German regions where high exit probabilities tend to occur in the small or medium size classes. The systematic picture suggests that this is not a 'failure' of the optimisation routine but rather might point to certain structural developments. In order to detect those, the further analysis is concentrated on two German regions: Northrhine-Westfalia (NRW) and Brandenburg which is also a test region in SEAMLESS. NRW is located in the very West of Germany, whereas Brandenburg has a border with Poland and was part of the former German Democratic Republic (GDR). Hence, differences in the agricultural structure of both regions are not only due to climatic differences but to historical reasons as well.

To begin with, the total farm numbers for NRW and Brandenburg are displayed in Figure 1 and Figure 2, respectively. The values until 2003 are the represented farm numbers taken from FADN, the values from 2003 to 2020 are the forecasted farm numbers.

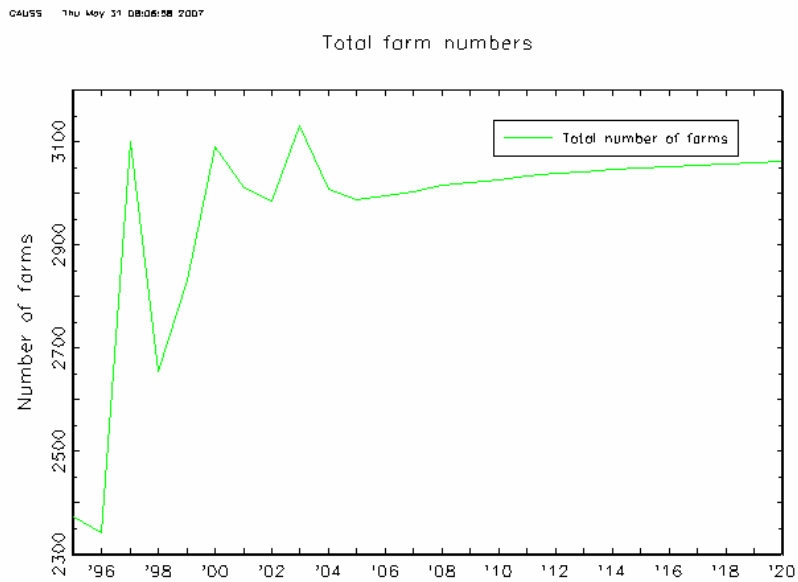
Figure 1: Total farm numbers NRW



Source: FADN, estimated.

The total number of farms in NRW was already halved in the observation period 1990 to 2003 and it is predicted that in 2020 only about 20% of the original amount of farms in 1990 will exist.

Figure 2: Total farm numbers Brandenburg



Source: FADN, estimated.

The time series for Brandenburg is relatively unstable, but the total number of farms seems to have increased during the observation period from 1995 to 2003. For the time after 2003 a slump of the total number of farms is predicted with slightly increasing farm numbers afterwards. From the figures it becomes apparent that the structural developments in both regions are, in fact, totally different from each other. Decreasing farm numbers in West Germany and increasing farm numbers in the Eastern part of Germany have been expected and originate in the different historical background, with the agricultural sector in NRW still being family farm based and a large farms dominated agriculture in Brandenburg. Thus, in Brandenburg only about 14% of the number of farms in NRW can be found, whereas the UAA of Brandenburg is about 80% the UAA of NRW. However, after the breakdown of the Iron Curtain and the German reunification some of the LPGs (the collectivised farms) might have split up into smaller parts which contribute to the increasing farm numbers, a typical pattern for many Eastern European countries as well.

Which are the most frequent farm types in the both regions is shown in Table 6 and Table 7 where the farm type shares are given for NRW and Brandenburg, respectively. The first column shows the shares in 1990 (1995 for East German regions), the beginning of the observation period, the second column represents the last year of the observation period, 2003. The third and fourth column show the predicted shares for 2013 and 2020.

Table 6: Farm type shares, NRW

	1990	2003	2013	2020
ARAB_S	5.64	0.00	0.00	0.00
ARAB_M	7.79	9.23	8.07	7.02
ARAB_L	6.98	13.69	12.14	10.63
SHGM_S	0.00	0.00	0.00	0.00
SHGM_M	0.37	0.00	0.00	0.00
SHGM_L	0.00	0.00	0.00	0.00
PERM_S	0.11	0.00	0.01	0.00
PERM_M	0.00	0.00	0.00	0.00
PERM_L	0.95	0.74	0.20	0.08
DARY_S	5.33	0.00	0.13	0.13
DARY_M	13.21	5.24	2.51	1.49
DARY_L	9.05	15.38	12.87	10.01
BEEF_S	1.30	0.00	1.05	1.05
BEEF_M	2.05	3.12	2.40	2.21
BEEF_L	1.36	4.12	7.08	9.52
PIGS_S	0.00	0.00	0.14	0.10
PIGS_M	1.45	2.17	1.51	1.12
PIGS_L	1.34	7.51	12.88	17.93
POLT_S	0.00	0.00	0.00	0.00
POLT_M	0.00	0.00	0.00	0.00
POLT_L	0.00	0.29	0.00	0.00
MIXF_S	4.51	0.00	1.61	1.45
MIXF_M	8.87	7.50	5.73	4.83
MIXF_L	8.41	16.20	12.94	11.24
MIXL_S	2.49	0.00	0.92	1.18
MIXL_M	7.18	2.34	2.98	3.65
MIXL_L	5.28	6.98	8.83	10.40
HORT_S	0.00	0.00	0.20	0.20
HORT_M	2.09	1.56	1.51	1.48
HORT_L	4.25	3.94	4.29	4.27

Source: FADN, estimated.

Table 7: Farm type shares, Brandenburg

	1995	2003	2013	2020
ARAB_S	0.00	0.00	0.43	0.33
ARAB_M	9.70	9.58	11.83	12.74
ARAB_L	21.09	33.22	40.60	43.73
SHGM_S	0.00	0.00	0.10	0.07
SHGM_M	4.77	2.97	2.20	1.67
SHGM_L	1.48	1.60	0.85	0.62
PERM_S	0.00	0.00	0.03	0.03
PERM_M	0.00	0.00	0.03	0.03
PERM_L	2.11	2.87	0.20	0.16
DARY_S	0.00	0.00	0.00	0.00
DARY_M	1.69	0.64	0.13	0.10
DARY_L	12.23	8.30	7.03	6.07
BEEF_S	4.22	0.00	0.20	0.16
BEEF_M	1.60	0.86	0.10	0.10
BEEF_L	3.16	3.51	2.63	2.12
PIGS_S	0.00	0.00	0.00	0.00
PIGS_M	0.00	0.00	0.00	0.00
PIGS_L	0.00	2.87	3.78	4.21
POLT_S	0.00	0.00	0.00	0.00
POLT_M	0.00	0.00	0.00	0.00
POLT_L	0.00	0.00	0.00	0.00
MIXF_S	5.90	0.00	0.00	0.00
MIXF_M	5.90	4.47	2.93	2.38
MIXF_L	23.20	18.21	14.43	11.85
MIXL_S	0.00	0.00	0.00	0.00
MIXL_M	0.00	0.64	0.00	0.00
MIXL_L	0.00	0.67	0.10	0.07
HORT_S	0.00	0.00	0.13	0.10
HORT_M	0.00	3.83	4.87	5.29
HORT_L	2.95	5.75	7.40	8.16

Source: FADN, estimated.

The farm types with the highest shares of farms are arable, dairy and mixed farms in both regions. In NRW also many hog and mixed livestock farms can be found or at least relatively high shares of those are predicted for the future. In Table 8 the three most frequent farm types at the four points in time for NRW and Brandenburg are displayed.

Table 8: Most frequent farm types NRW and Brandenburg

	NRW			Brandenburg		
	1.	2.	3.	1.	2.	3.
1990/1995	DARY_M	DARY_L	MIXF_M	ARAB_L	MIXF_L	DARY_L
2003	MIXF_L	DARY_L	ARAB_L	ARAB_L	MIXF_L	ARAB_M
2013	MIXF_L	PIGS_L	DARY_L	ARAB_L	MIXF_L	ARAB_M
2020	PIGS_L	MIXF_L	DARY_L	ARAB_L	ARAB_M	MIXF_L

Source: FADN, estimated.

Whereas there is much movement in the most frequent farm types in NRW, the most frequent farm types in Brandenburg are relatively stable. It is characteristic for the structural developments in West Germany that there have been two medium sized farm types among the three most frequent farm types in NRW in 1990, whereas already in 2003 only large farm sizes can be found among the first three farm types. As regards specialization it is predicted that livestock keeping farm types will gain more importance in NRW. In Brandenburg there seems to be a strong development in favour of arable farms. It is predicted that the already very high share of large arable farms in 2003 (33%) will increase to nearly 44% in 2020. And according to the prediction the second highest share in 2020 will come from medium arable farms (ca. 13%). It is quite surprisingly and contrary to the results for NRW that there have been no medium sized farm types among the three most frequent farm types in 1995, the first observation year, but medium arable farms appeared as third important farm type in 2003 and will even gain importance until 2020.

From Table 6 and Table 7 it is obvious that the distribution of farm types is much more balanced in NRW than in Brandenburg where, as already seen above, arable farms have by far the highest contribution in terms of farm numbers. In Table 9 the number of farm types with a share of more or equal to 5, respectively 10% are shown for each region.

Table 9: Farm type shares

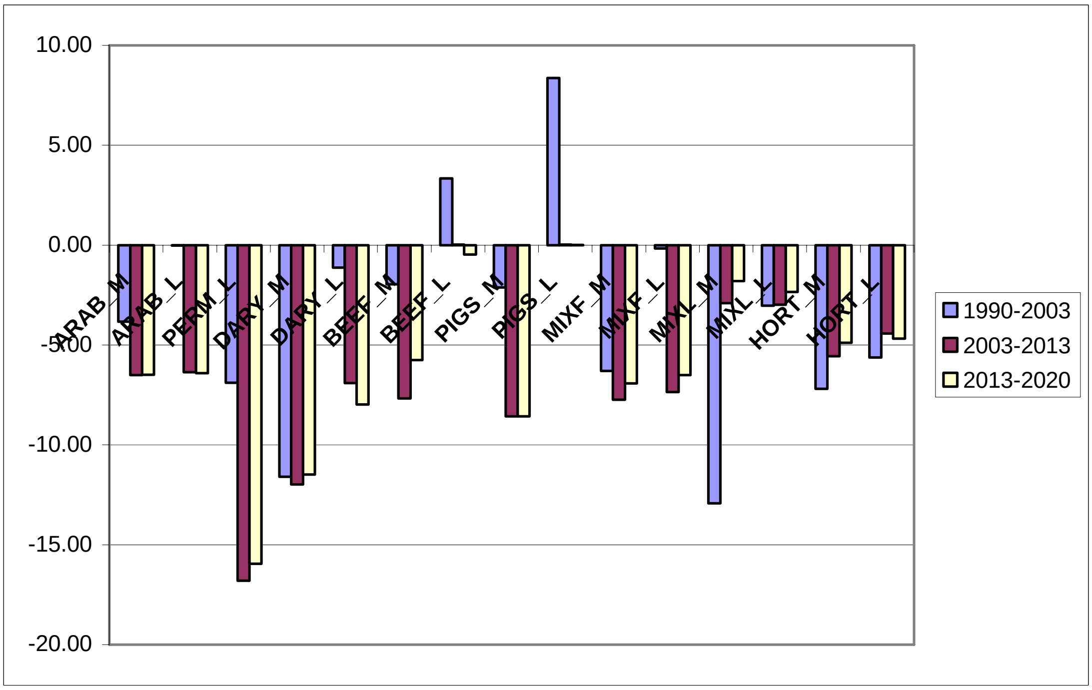
	NRW, 5%				Brandenburg, 5%			
	1990	2003	2013	2020	1990	2003	2013	2020
small	2	0	0	0	1	0	1	0
medium	4	3	2	1	2	1	1	2
large	4	5	6	6	3	4	4	4
sum	10	8	8	7	6	5	6	6
	NRW, 10%				Brandenburg, 10%			
small	0	0	0	0	0	0	0	0
medium	1	0	0	0	0	0	1	1
large	0	3	4	5	3	2	2	2
sum	1	3	4	5	3	2	3	3

Source: FADN, estimated.

For NRW it is obvious that the number of farm types with a more than 5% share shrinks, whereas the number of farm types with more than 10% increases. This could be interpreted as a concentration towards a few, perhaps more efficient farm types. The number of farm types for both, the 5% and the 10% category in Brandenburg, in contrast, remains nearly the same throughout the observation and forecasted period. Together with the observations made in context of Table 8 and the figures on the total farm numbers, it might be followed that there is much more fluctuation in NRW than in Brandenburg indicating that the structural change in Brandenburg is nearly completed, whereas in NRW still considerable changes take place.

The figures below (Figure 3 and Figure 4) show the growth rates of the single farm types for the time periods 1990-2003 (1995-2003 for Brandenburg), 2003-2013 and 2013-2020. Due to zero values in the data the growth rates could not be calculated for all farm types but are given only for those where for all three time periods data was available.

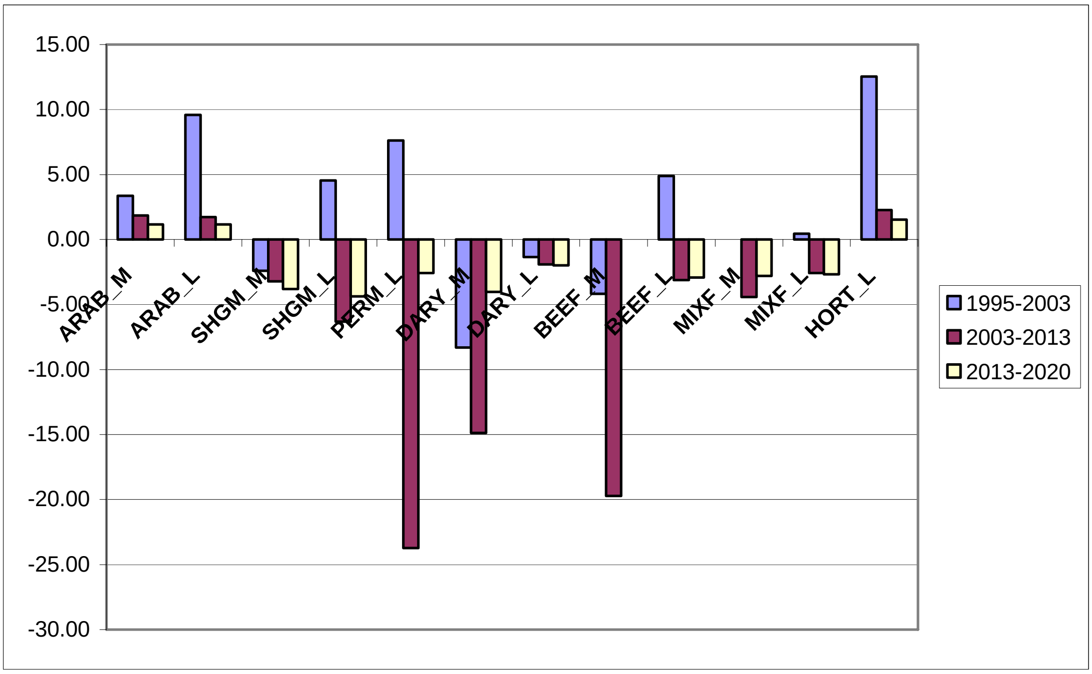
Figure 3: Growth rates NRW



Source: FADN, estimated.

In NRW positive growth rates can only be found for the time period 1990-2003 for large beef and large pig farms. However, for both farm types the speed of growth slows down for the forecasted periods and is near to zero for large beef farms from 2003-2013 and for large hog farms from 2003-2020. For large beef farms the growth rate is predicted to be negative for the time period 2013-2020. Especially high negative growth rates are predicted for medium and large dairy farms. The overall growth rate for 1990-2003 is -5.08, -5.23 for 2003-2013 and -4.61 for 2013-2020.

Figure 4: Growth rates Brandenburg



Source: FADN, estimated.

There are quite a number of farm types in Brandenburg for which positive growth rates are reported although the period of growth lay mostly in the past (1995-2003) and considerably less positive growth rates are predicted for the future. Farm types with positive growth rates from 1995-2003 and negative growth rates afterwards are large sheep and goat farms, large permanent farms, large beef farms and large mixed farms. Positive growth rates during all time periods can be found for medium and large arable farms and large horticultural farms. The overall growth rates are 3.54 for 1995-2003, -0.29 for 2003-2013 and 0.09 for 2013-2020. Particularly high negative growth rates are predicted for large permanent crops farms, medium dairy and medium beef farms, in each case for the time period 2003-2013.

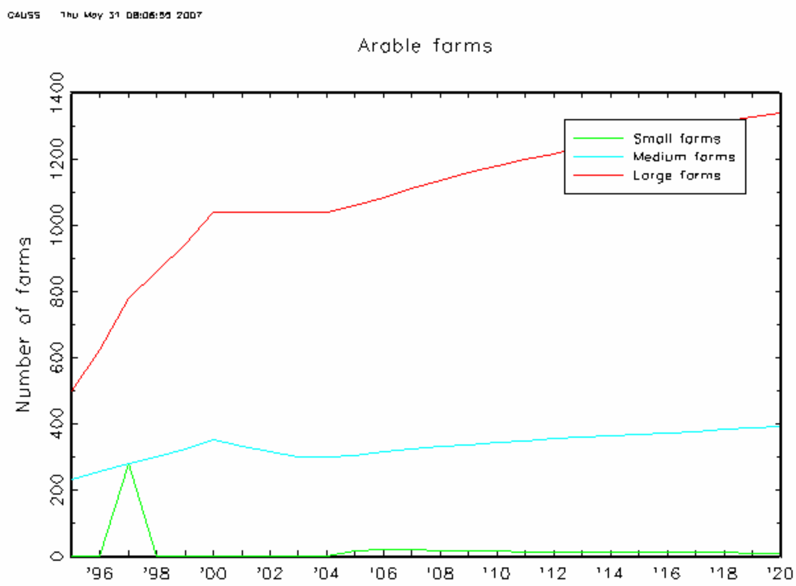
Next, figures for the most important specialisation classes for each region are shown and compared to each other. According to the preceding analysis these are arable, dairy, and mixed farms. For NRW also the development of the number of pig farms is shown.

Figure 5: Arable farms NRW

Source: Estimated.

The relatively considerable amount of small arable farms in NRW in 1990 seems to have disappeared altogether in the time period between 1997 and 2000. The amount of medium sized farms fell in the beginning to the mid-nineties but then increased again until 2000. The increase was followed by a sharp decline in the three-years period from 2000 to 2003 such that the overall growth rate from 1990 to 2003 is negative (Figure 3). The number of large arable farms increased significantly during the nineties and fell abruptly from 2000 to 2003. The transition probabilities (Table 10) suggest that the small farms have not left the sector but rather increased in size. The medium farms have high probabilities to stay in the same farm type or alternatively are supposed to grow to a large farm. From the arable farms only large farms are allowed to leave the sector with a probability of 0.107. There are no entries predicted for NRW. Whether it makes sense that small and medium farms have an exit probability of zero is rather questionable. However, looking at the pace of the decline of the amount of large farms in the last observation period 2000/2003 it is reasonable that their probability to exit is larger than for the medium sized farms. One should also keep in mind that the decline of the number of large farms does not naturally mean that the cultivated area or production of arable crops is shrinking as well, but is most likely due to a number of large farms which are still growing in size urging others to leave the sector. However, such a development could only be detected by introducing a fourth 'very large farms' size class.

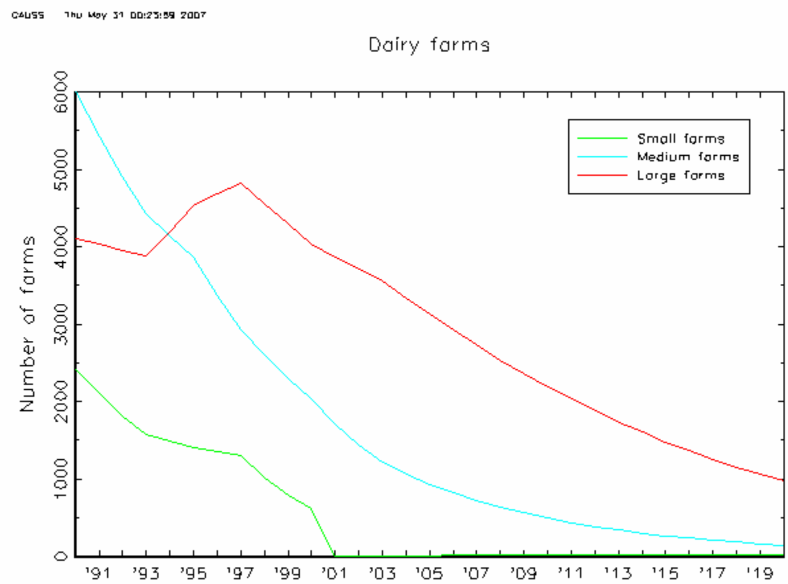
Figure 6: Arable farms Brandenburg



Source: Estimated.

Another picture is drawn for arable farms in Brandenburg. Whereas here nearly no small farms exist, does the number of medium and large arable farms increase. The probability estimates are a bit inconsistent regarding the small size category. There is a relatively high probability value for entry but no farms actually stay in the small size category but rather grow to medium and large farms or, most likely, exit the sector making the small size category a transition farm type. However, this is a quite typical pattern if only very few farms can be observed in a farm type indicating that the solver has problems to find the 'real' transition probabilities. Obviously the predicted number of farms is not much affected by this behaviour. A solution to the problem could be to constrain the small size category probabilities to zero altogether when it comes to the fine tuning of the TPMs for selected regions. Medium arable farms have a relatively high probability to become a large farm and a very small probability to leave the sector. Large farms have a very high probability to stay a large farm. There is a small probability for them to change to the medium size class, the exit probability is equal to zero. Large arable farms in Brandenburg had with about 21% a very high share of total farms already at the beginning of the observation period in 1995. It is predicted that the share of large arable farms will lie at about 44% in 2020.

Figure 7: Dairy farms NRW

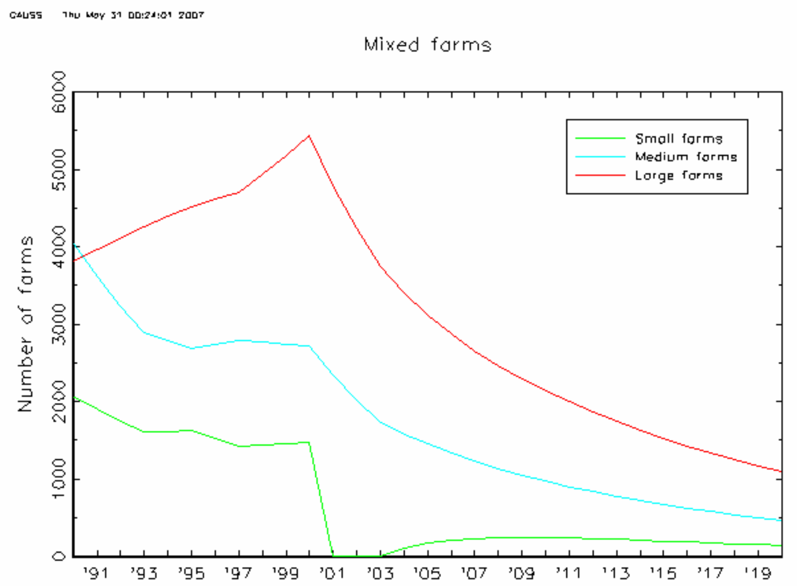


Source: Estimated.

The figure for the dairy specialisation class in NRW shows a characteristic pattern of agricultural structural change in West Germany: small farms disappear sometime during the observation period, i.e. most likely at the end of the nineties or at the beginning of the 21st century and the number of large farms exceeds the number of medium farms in the period between the beginning of the observation period and the mid-nineties. Typical is also that the number of large farms increases until the end of the nineties and then decreases relative abruptly. According to the TPM do the small dairy farms either stay in the smallest size class or become medium sized farms. The medium farms have a relatively high probability to change to the large size category and a very small probability to shrink in size and become a small farm. As already seen for the arable farms do only the large farms leave the sector via the exit class. The probability of exit is thereby quite high with 0.1. Again, this pattern of exit probabilities is rather questionable and should be taken care of in future model versions.

The figure for Brandenburg dairy farms is not shown since there are nearly no dairy farms in Brandenburg and the 15 sample farms rule might be violated otherwise. However, as can be seen from the shares in Table 7 there are no small and only very few medium dairy farms. And the growth rates in Figure 4 indicate that also the number of large dairy farms is obliged to decrease. According to the TPM there are no entries to the small and medium size classes and the value for the exit probability is one, meaning that all farms that could possibly enter these classes will directly move out of the sector again. This holds for example for the small amount of large dairy farms that is supposed to move to the medium size class with a probability of 0.012. This could be interpreted as a failure of the estimation procedure as well since it has no other meaning than the exit category itself. However, the values displayed represent the optimal solution in technical terms and corrections regarding consistency could be carried out only by imposing additional constraints or making use of *a priori* information. The probability of exit for large dairy farms is denoted with 0.009 and consequently the probability to stay a large farm is very high with 0.979.

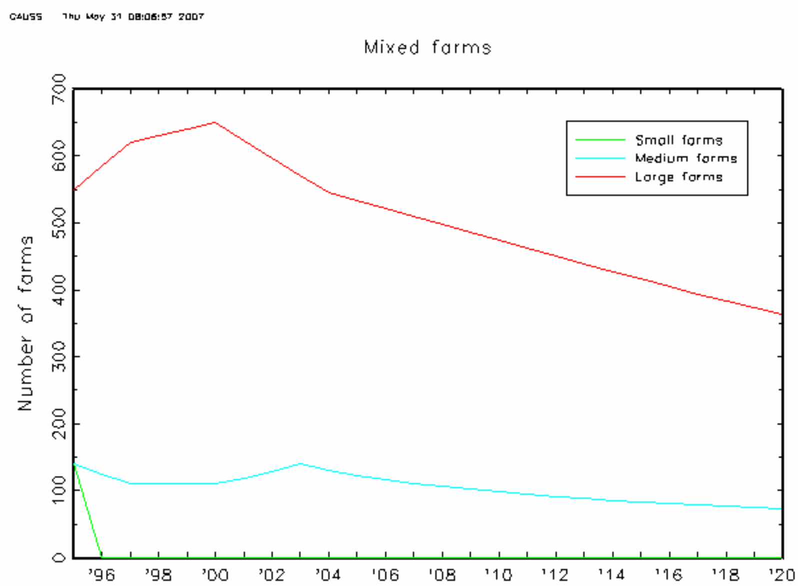
Figure 8: Mixed farms NRW



Source: Estimated.

The structural development pattern for mixed farms in NRW is nearly the same as seen already for the arable farms and described for the dairy farm specialisation, although there are slightly more small farms predicted than for the other both specialisation classes. Small farms have a relatively low probability to stay in the small size category but rather change to the medium or large size class. According to the transition probabilities do medium farms change to the small as well as to the large size category. As already seen in the other specialisation types do only large farms leave the sector via the exit category. Large farms have also a small probability to become a medium sized farm.

Figure 9: Mixed farms Brandenburg

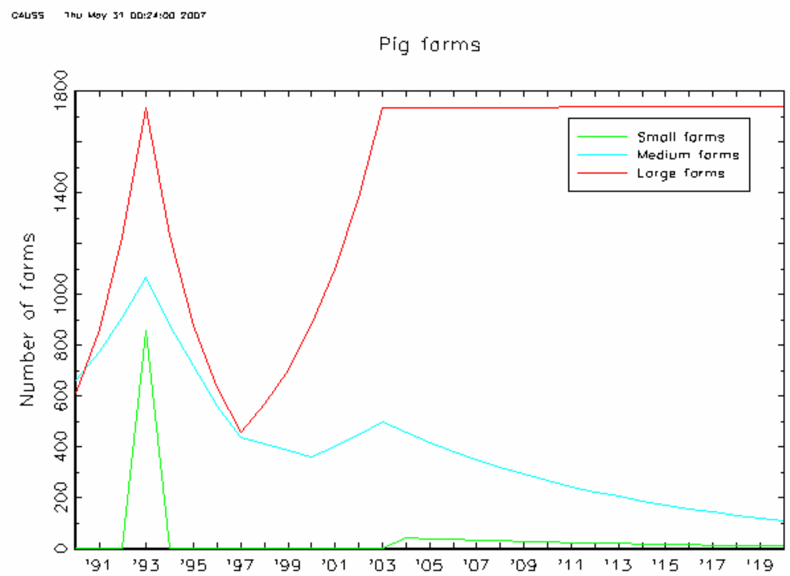


Source: Estimated.

In Brandenburg there are by far more large mixed farms than farms in the other size categories from which especially the small size category is only marginal. Accordingly, is the probability to exit from the small size class equal to one. Medium farms have a relatively small probability to stay in the medium size class and a high probability to leave the sector. Large farms are most likely to stay in the same size class. There is a small probability to change to the medium size class and a slightly higher probability for entries to the large size class.

For NRW also the pig specialisation type is shown since it significantly contributes to the total share of farms in the region (Figure 10).

Figure 10: Pig farms NRW



Source: Estimated.

The heavy fluctuations that are shown in Figure 10 for large hog farms quite often appear in the FADN time series and make it difficult to estimate the 'correct' transition probabilities or even judge in which direction they should point. Whereas the time series for medium pig farms is still relatively even and it is clear that the number of medium sized farms tends to decrease, there are really high values for large pig farms in one and more than 50% less farms in the other survey year. In this case, the solver finally chose to tie in with the last observation point and a high amount of large pig farms is predicted which is kept relative constantly over the forecasted period. The TPM says that a small amount of farms moves from the medium to the small size class from where they exit the sector (exit probability 1.0). There is only a probability of 0.001 predicted to change from the medium to the large size class. However, farms in the large size class are predicted to stay there with a probability of one. As all other farm types are obliged to decrease in the long run (Figure 3) and only large pig farms are foreseen to have a very slight positive growth rate, it is no surprise that large pig farms are predicted to become the farm type with the highest share (18%) of farms in 2020. This is however a rather questionable result since there is an ongoing process of concentration in hog production in NRW and it might be necessary to adjust the transition probabilities for pig farms manually in order to come to more plausible conclusions.

Generally, there seem to be heavier changes in NRW, whereas the values for Brandenburg give the impression that structural change might be nearly completed. Whereas farm numbers are clearly decreasing in NRW, for Brandenburg constant or even slightly increasing farm numbers are predicted with positive probabilities for entry in some cases which could possibly result from the splitting of former collective farms. The lesson learnt is that all numbers shown need to be interpreted carefully and in dependence of each other. The effect that these results will have on the SEAMLESS model chain is not foreseeable and needs to be tested empirically.

Table 10: TPM Northrhine-Westfalia

	ARAB_S	ARAB_M	ARAB_L	SHGM_S	SHGM_M	SHGM_L	PERM_S	PERM_M	PERM_L	DARY_S	DARY_M	DARY_L	BEEF_S	BEEF_M	BEEF_L	PIGS_S	PIGS_M	PIGS_L	POLT_S	POLT_M	POLT_L	MIXF_S	MIXF_M	MIXF_L	MIXL_S	MIXL_M	MIXL_L	HORT_S	HORT_M	HORT_L	EXIT
ARAB_S	0.839	0.000	0.000																												0.000
ARAB_M	0.039	0.935	0.000																												0.000
ARAB_L	0.122	0.065	0.893																												0.000
SHGM_S				0.000	0.197	0.000																									0.000
SHGM_M				0.000	0.780	0.000																									0.000
SHGM_L				0.000	0.022	0.000																									0.000
PERM_S							0.000	0.000	0.030																						0.000
PERM_M							0.000	0.000	0.000																						0.000
PERM_L							0.001	0.000	0.833																						0.000
DARY_S										0.899	0.004	0.000																			0.000
DARY_M										0.101	0.878	0.000																			0.000
DARY_L										0.000	0.118	0.900																			0.000
BEEF_S													0.777	0.081	0.000																0.000
BEEF_M													0.098	0.897	0.000																0.000
BEEF_L													0.000	0.021	0.989																0.000
PIGS_S																0.000	0.085	0.000													0.000
PIGS_M																0.000	0.914	0.000													0.000
PIGS_L																0.000	0.001	1.000													0.000
POLT_S																			0.000	0.000	0.000										0.000
POLT_M																			0.000	0.000	0.000										0.000
POLT_L																			0.000	0.000	0.128										0.000
MIXF_S																						0.740	0.058	0.000							0.000
MIXF_M																						0.049	0.858	0.025							0.000
MIXF_L																						0.211	0.083	0.872							0.000
MIXL_S																									0.746	0.059	0.006				0.000
MIXL_M																									0.119	0.888	0.019				0.000
MIXL_L																									0.135	0.052	0.944				0.000
HORT_S																												0.432	0.069	0.000	0.000
HORT_M																												0.000	0.803	0.052	0.000
HORT_L																												0.000	0.129	0.909	0.000
EXIT	0.000	0.000	0.107	1.000	0.000	1.000	0.999	1.000	0.137	0.000	0.000	0.100	0.125	0.000	0.011	1.000	0.000	0.000	1.000	1.000	0.872	0.000	0.000	0.103	0.000	0.000	0.031	0.568	0.000	0.040	1.000

Source: Estimated.

R^2: 0.982.

Table 11: TPM Brandenburg

	ARAB_S	ARAB_M	ARAB_L	SHGM_S	SHGM_M	SHGM_L	PERM_S	PERM_M	PERM_L	DARY_S	DARY_M	DARY_L	BEEF_S	BEEF_M	BEEF_L	PIGS_S	PIGS_M	PIGS_L	POLT_S	POLT_M	POLT_L	MIXF_S	MIXF_M	MIXF_L	MIXL_S	MIXL_M	MIXL_L	HORT_S	HORT_M	HORT_L	EXIT
ARAB_S	0.000	0.000	0.000																												0.141
ARAB_M	0.073	0.850	0.043																												0.050
ARAB_L	0.209	0.146	0.957																												0.169
SHGM_S				0.000	0.000	0.000																									0.030
SHGM_M				0.308	0.767	0.340																									0.037
SHGM_L				0.000	0.094	0.660																									0.015
PERM_S							0.000	0.000	0.214																						0.000
PERM_M							0.000	0.000	0.147																						0.000
PERM_L							0.000	0.000	0.448																						0.035
DARY_S										0.000	0.000	0.000																			0.000
DARY_M										0.000	0.000	0.012																			0.010
DARY_L										0.000	0.000	0.979																			0.003
BEEF_S													0.000	0.700	0.000																0.042
BEEF_M													0.000	0.232	0.000																0.026
BEEF_L													0.000	0.000	0.969																0.000
PIGS_S																0.000	0.000	0.000													0.000
PIGS_M																0.000	0.000	0.000													0.000
PIGS_L																0.000	0.000	1.000													0.024
POLT_S																			0.000	0.000	0.000										0.000
POLT_M																			0.000	0.000	0.000										0.004
POLT_L																			0.000	0.000	0.000										0.000
MIXF_S																						0.000	0.000	0.000							0.000
MIXF_M																						0.000	0.764	0.042							0.000
MIXF_L																						0.000	0.000	0.958							0.077
MIXL_S																									0.000	0.000	0.000				0.000
MIXL_M																									0.000	0.000	0.000				0.000
MIXL_L																									0.000	0.000	0.459				0.014
HORT_S																												0.000	0.000	0.000	0.040
HORT_M																												0.228	0.629	0.235	0.039
HORT_L																												0.000	0.371	0.765	0.022
EXIT	0.718	0.004	0.000	0.692	0.139	0.000	1.000	1.000	0.191	1.000	1.000	0.009	1.000	0.068	0.031	1.000	1.000	0.000	1.000	1.000	1.000	1.000	0.236	0.000	1.000	1.000	0.541	0.772	0.000	0.000	0.221

Source: Estimated.

R^2: 0.972.

6 Link to SEAMLESS-IF

There are basically two applications for the structural change module. One is its use in the model chain and the other one is its role in post-model analysis concerning social indicators for WP2. Here, the technical realisation and potential effects of the integration into the model chain are discussed.

6.1 Technical realisation

Theoretically, only the transition probabilities will enter SEAMLESS-IF. They represent the main outcome of the structural change module from which all other predicted values (e.g. farm type shares, growth rates) can be derived. It is intended that the transition probabilities will be established in the SEAMLESS database from where they can be called and further processed in order to meet the demands of different applications of which the most important one might be the usage in the up-scaling procedure from the farm to the market level in EXPAMOD. In EXPAMOD the supplied quantity of various products estimated by the FSSIM models will be weighted by the respective farm type shares for each region to achieve regional coverage. For that purpose a GAMS code will be written which uses the transition probabilities together with the respective actual base year data on the number of farms (both stored in the SEAMLESS database) to calculate forecasts on the farm numbers (equation (0.5)) from which then the farm type shares can be derived. The shares are equal to the weighting factors in EXPAMOD (Bezlepkina et al. 2006).

However, the envisaged procedure described above might not work because the stored data on the number of farms is obliged to FADN confidentiality rules. Data that is based on less than 15 sample farms is not shown such that quite a lot of farm types are in danger of getting lost for the analysis. However, as the FSSIM models are also based on the limited data there might even be no need for weighting factors for these farm types. The current idea is to build up a ‘remaining farms’ farm type which then could be extrapolated with an aggregate of the individual aggregation weights. There is an ongoing discussion on this issue. Until the confidentiality problem is solved and the SEAMLESS database is fully operable, the farm type shares are directly calculated in the structural change module and the data exchange with EXPAMOD is carried out via Excel-files.

Still another complication occurs when non-stationary transition probabilities become available because the exogenous variable values are likely to change during the simulation period. However, the non-stationary transition probabilities can be recalculated for each transition period depending on forecasts on the significant explanatory variables (the forecasts might be a result of an earlier model run, see chapter 6.2). According to equation (0.19), then, only the priors and shadow prices for the constraints should be stored in the database since all other factors are variable.

Furthermore, it is possible to re-estimate the (stationary and non-stationary) transition probabilities whenever new FADN data become available. This should be done in order to capture not only effects from a changing environment regarding explanatory variables as described above but also account for potential changes in the pattern of structural change itself which might have not been predicted by the estimations based on older FADN data. It is proposed to re-estimate the transition probabilities on a regular basis (e.g. every 4-5 years) or after the occurrence of major events such as policy shocks or the EU accession of large agricultural producers like Ukraine or Turkey. However, updating the transition probabilities is not straight forward and should therefore be done by the research group formed for maintaining the system once the development phase of SEAMLESS is completed.

6.2 How structural change affects the FSSIM-EXPAMOD-SEAMCAP model chain

6.2.1 The use of stationary transition probabilities

As explained above, structural change measured with transition probabilities has a direct impact in the up-scaling process in EXPAMOD. Let us take the example of one region with only two farm types to illustrate this. One farm type is specialised in cereals production, the other one in cattle production. Both farms produce cattle and cereals. We further assume that the specialisation degree is correlated with production costs meaning that it is cheaper to produce cereals on a farm specialised in cereals production and vice versa. Let us assume that EXPAMOD estimates for the two farms the following own price supply elasticities:

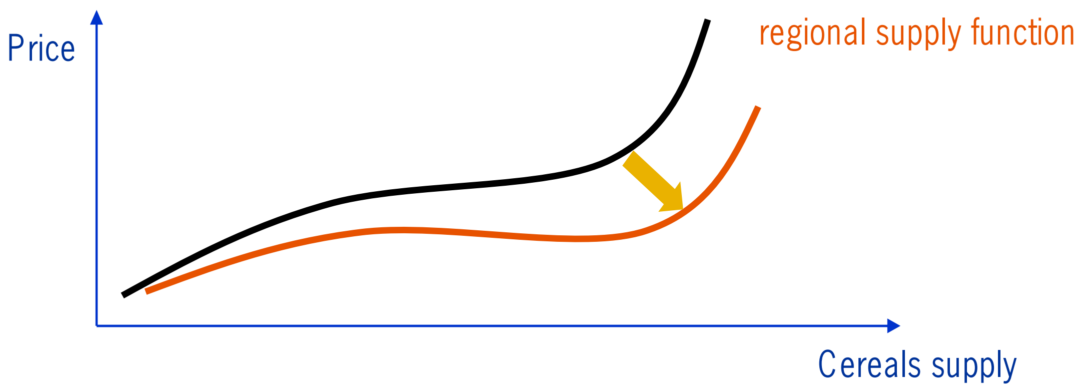
	Farm 1	**Farms 2**
Cereals	**1.3**	**0.9**
Cattle	**0.8**	**1.2**

The regional own price supply elasticity is a weighted average of the elasticities per farm. The weighting factors, as explained above, are the farm type shares. If we assume that both farms have a weight of 50% in this particular region, we derive regional elasticities of 1.1 for cereals and 1 for cattle.

Let us now include structural change into this procedure. If the results of the Markov chain analysis are predicting that we have in every year a certain transition probability on cattle farms to convert into a cereals farm so that the weights in the simulation year are now 75% for cereals farms and 25% for cattle farms, the new regional supply elasticities would be 1.2 for cereals and 0.9 for cattle.

The supply elasticities in turn impact on the regional supply behaviour in SEAMCAP, since they determine the regional marginal cost curves driving supply what is simplified shown for the cereals example in Figure 11.

Figure 11: Example of a regional supply curve



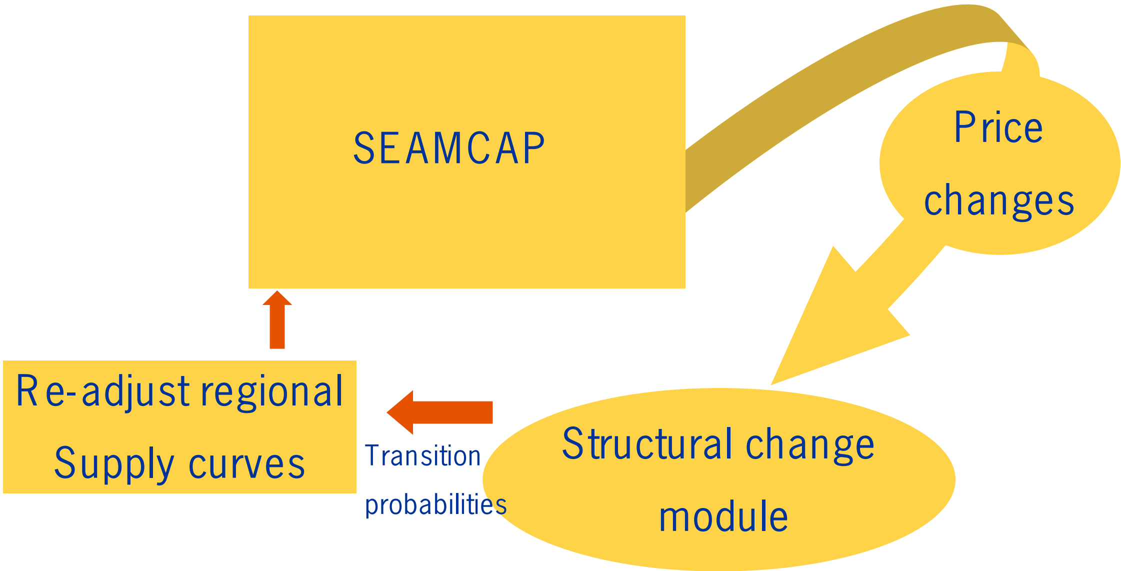
The regional supply curve without including transition probabilities (black) would be steeper than the one including them (red).

Currently the supply elasticities coming from EXPAMOD are only used as prior information in the calibration process of the regional supply functions in SEAMCAP. The resulting supply elasticities used in SEAMCAP are not exactly the same. On the one hand a complete congruency is not desirable because SEAMCAP uses some other restrictions than the FSSIM models, on the other hand the current calibration method in SEAMCAP is not flexible enough to allow for a stronger convergence between the two models. This issue will be addressed in prototype 3 of SEAMLESS-IF.

6.2.2 The potential use of non-stationary transition probabilities

Currently it is not clear whether the estimation of non-stationary transition probabilities will identify significant impact of macroeconomic variables on transition probabilities of farm types. But if some variables that are endogenous to SEAMCAP turn out to be significant the following procedure could be applied. If we assume for example an aggregated price index of agricultural commodities has an impact on the transition of one farm type to another, the regional supply curves in SEAMCAP will change within one policy simulation in an iterative process.

Figure 12: Non-stationary transition probabilities in SEAMCAP



Principally SEAMCAP receives a policy shock that leads to a new set of prices of agricultural goods. Those price changes would then be used in the structural change module to calculate the new regional distribution of farm types (either by adjustment of the existing transition probabilities via equation (0.19) or by re-estimating the transition probabilities). Using them and the supply elasticities per farm types, the supply curves can be readjusted as described in Figure 11 and SEAMCAP can be started for another round.

7 Conclusions

After a brief repetition of some theoretical considerations regarding structural change the farm typology applied and the data available for the analysis are explained. The main part of the deliverable comprises the description of the estimation technique. Different methodologies for stationary and non-stationary Markov chain approaches are presented. So far only a simplified model is estimated with the more advanced model versions being under development. With regard to methodology two challenges are to be solved in follow-ups for this deliverable: firstly, the amount of parameters estimated should be increased to the estimation of transition probabilities for transitions between specialisation classes as well. For this a thorough analysis of the micro data is envisaged such that information on transitions that have occurred in the single farm data time series could help in the estimation of the enlarged model. Secondly, a non-stationary model needs to be developed which allows the estimation of time-varying transition probabilities in dependence of changes in certain explanatory variables. Potential explanatory variables for structural change are listed in the very beginning of the deliverable, but the exact set of variables which is able to explain structural change across the different specialisations still needs to be defined. Results are shown for the reduced model and analysed and compared to each other in detail for two German regions. The main findings are that structural change is an ongoing process which might considerably differ even between regions of the same country. As the transition probabilities are estimated rather ‘freehand’ without imposing specific assumptions regarding the underlying processes as done in most other studies some inconsistencies in the resulting transition probability matrices can be observed. But in general few problems occurred when applying these probabilities for the prediction of future farm numbers. The last chapter deals with the integration of the structural change module into SEAMLESS-IF. Regarding the technical realisation it is argued that the data exchange with EXPAMOD should be carried out manually with specifically prepared files until the SEAMLESS data base is fully operable and it is clear for which farm types information will be provided by the FSSIM models. Theoretical considerations are given on how other parts of the model chain could be affected by the farm type shares derived from the structural change module.

The next update of this deliverable is due in month 39.

References

Ahearn, M., Yee, J., Huffman, W., 2002. R&D, Productivity, and Structural Change in U.S. Agriculture, 1960-1996. National Bureau of Economic Research Summer Institute, Conference on Research in Income and Wealth, Cambridge, Massachusetts.

Andersen, E., Verhoog, A., Elbersen, B., Godeschalk, F., Koole, B., 2006. A multidimension farming system typology.

Balmann, A., 1997. Farm Based Modelling of Regional Structural Change. European Review of Agricultural Economics 25 (1), 85-108.

Bezlepkina, I., Pérez, I., Heckelei, T., Romstad, E., Oude Lansink, A., 2006. Component to Statistically Extrapolate from FSSIM Models to Other Farm Types and Regions Including Aggregation to NUTS2: Motivation, Description and Prototype. SEAMLESS Integrated Project, EU 6th Framework Programme, contract no. 010036-2.

Boehlje, M., 1990. Alternative Models of Structure Change in Agriculture and Related Industries. University of Minnesota, Department of Agricultural and Applied Economics, Staff Paper Series P40-41.

Bremmer, J., Oude Lansink, A.G., Olson, K.D., Baltussen, W.H., Huirne, R.B., 2004. Analysis of Farm Development in Dutch Agriculture and Horticulture. Strategic Decision Making and Firm Development in Dutch Horticulture 23-39.

Cochrane, W., 1958. Farm Prices: Myth and Reality. University of Minnesota Press, Minneapolis.

European Commission, 2005. FADN.

European Commission, 2006. Structure of agricultural holdings.

European Commission, 2000. Farm Structure. Luxembourg.

Goddard, E., Weersink, A., Chen, K., Turvey, C.G., 1993. Economics of Structural Change in Agriculture. Canadian Journal of Agricultural Economics 41 475-86.

Golan, A., Judge, G., Miller, D., 1996. Maximum Entropy Econometrics. John Wiley and Sons, Chichester.

Golan, A., Vogel, S., 2000. Estimation of non-stationary social accounting matrix coefficients with supply side information. Economic Systems Research 12 447-471.

Hallam, A., 1991. Economies of Size and Scale in Agriculture: An Interpretative Review of Empirical Measurement. Review of Agricultural Economics 13 (1), 155-172.

Hallam, A., 1993. Empirical Studies of Size, Structure, and Efficiency in Agriculture. In: Hallam, A. (Eds.), Size, Structure, and the Changing Face of American Agriculture. Westview Press, Boulder, Colorado, pp. 204-231.

Harrington, D.H., Reinsel, R.D., 1995. A Synthesis of Forces Driving Structural Change. Canadian Journal of Agricultural Economics Special Issue (43), 3-14.

Jongeneel, R., Longworth, N., Huettel, S., 2005. Dairy Farm Size Distribution in East and West: Evolution and Sensitivity to Structural and Policy Variables: Case-Studies of the Netherlands, Germany, Poland, and Hungary. XIth International Congress of the EAAE, Copenhagen, Denmark.

Karantininis, K., 2002. Information-based estimators for the non-stationary transition probability matrix: an application to the Danish pork industry. Journal of Econometrics 107 275-290.

Kim, C., Schluter, G., Schaible, G.M.A., Hallahan, C., 2005. A Decomposed Negative Binomial Model of Structural Change: A Theoretical and Empirical Application to U.S. Agriculture. Canadian Journal of Agricultural Economics 53 161-176.

Knutson, R.D., Penn, J., Boehm, W.T., 1990. Agricultural and Food Policy, 2nd Edition. Prentice Hall.

Lee, T.C., Judge, G.G., 1996. Entropy and Cross-Entropy Procedures for Estimating Transition Probabilities from Aggregate Data. In: Berry, D.A., Chaloner, K.M., Geweke, J.K. (Eds.), Bayesian Analysis in Statistics and Econometrics - Essays in Honor of Arnold Zellner. John Wiley & Sons, Inc., New York, pp. 153-164.

MacRae, E.C., 1977. Estimation of Time-Varying Markov Processes with Aggregate Data. Econometrica 45 (1), 183-198.

Oehmke, J., Schimmelpfennig, D.E., 2004. Quantifiying Structural Change in U.S. Agriculture: The Case of Research and Productivity. Journal of Productivity Analysis 21 297-315.

Penn, J., 1979. The Structure of Agriculture: An Overview of the Issue, Structural Issues of American Agriculture, ESCS, USDA, Report 438.

Stanton, B.F., 1993. Farm Structure: Concept and Definition. In: Hallam, A. (Eds.), Size, Structure, and the Changing Face of American Agriculture. Westview Press, Boulder, Colorado, pp. 14-29.

Stavins, R.N., Stanton; B. F., B. F., 1980. Using Markov Models to Predict the Size Distribution of Dairy Farms, New York State, 1968-1985. Cornell University, Department of Agricultural Economics.

Tonini, A., 2007. Agriculture and Dairy in Eastern Europe after Transition focused on Poland and Hungary, Wageningen University, Wageningen, Netherlands.

Tweeten, L., 1984. Causes and Consequences of Structural Change in the Farming Industry, FAC Report No. 207. Washington, D.C.: National Planning Association.

U.S. Congress, O.o.T.A., 1985. Factors Contributing to Structural Change in Agriculture. In: U.S. Congress, O.o.T.A. (Eds.), Technology, Public Policy, and the Changing Structure of American Agriculture: A Special Report for the 1985 Farm Bill Technology, Public Policy, and the Changing Structure of American Agriculture: A Special Report for the 1985 Farm Bill . Washington, DC, pp. 107-118.

Zepeda, L., 1995a. Technical Change and the Structure of Production: A non-stationary Markov Analysis. European Review of Agricultural Economics 22 41-60.

Zepeda, L., 1995b. Asymmetry and Nonstationarity in the Farm Size Distribution of Wisconsin Milk Producers: An Aggregate Analysis. American Journal of Agricultural Economics 77 (2), 837-852.

Zimmermann, A., Heckelei, T., Pérez, I., 2006. Working Paper - Literature Review of Approaches to Estimate Structural Change.

Glossary

Structural change	Change of the number of farms in certain farm types.
Transition probability	Probability for a farm to change from one farm type to another in a specific time period.

Appendices

Table 12: Economic size thresholds applied by the Commission (in ESU)

	1990	1993	1995	1997	2000	2001	2002	2004
Belgium	12	12	12	12	16	16	16	16
Denmark	4	8	8	8	8	8	8	8
Germany	8	8	8	8	8	8	8	8
Greece	2	2	2	2	2	2	2	2
Spain	2	2	2	2	2	2	2	2
France	8	8	8	8	8	8	8	8
Ireland	2	2	2	2	2	2	2	2
Italy	2	2	2	2	2	2	4	4
Luxembourg	8	8	8	8	8	8	8	8
The Netherlands	16	16	16	16	16	16	16	16
Austria		8	8	8	8	8	8	8
Portugal	1	1	1	1	1	2	2	2
Finland		8	8	8	8	8	8	8
Sweden		8	8	8	8	8	8	8
United Kingdom (Northern Ireland)	4	4	4	4	4	8	8	8
United Kingdom	8	8	8	8	8	16	16	16
Hungary								2
Latvia								2
Lithuania								2
Estonia								2
Czech Republic								4
Poland								2
Slovakia								6
Slovenia								2
Malta								8
Cyprus								1
Bulgaria								2
Romania								2

Source: European Commission 2005.