Structural convergence between the dairy sectors

of the EU-27 Member States since

the Eastern Enlargement

Winand de Jong, Rico Ihle, Roel Jongeneel

Agricultural Economics and Rural Policy (AEP) group, Wageningen University

Corresponding author: winand.dejong@wur.nl

Paper prepared for presentation for the 162nd EAAE Seminar

The evaluation of new CAP instruments: Lessons learned and the road ahead

April 26-27, 2018

Corvinus University of Budapest

Budapest, Hungary

Abstract

Cohesion is one of the core goals of European integration. In 2004 and 2007, the EU welcomed twelve new members which added their agricultural sectors to the EU's single market. Especially their dairy sectors showed substantial competitive disparities to old Member states. We therefore examine structural convergence in farm-gate milk prices, productivity and farm income across EU's regional dairy sectors since 2004. We find that price dispersion decreased from 2004-2007. Kernel density plots and Markov chain estimates show that there was a high probability for regions to stay in the lower income- and productivity classes.

Keywords: convergence, dairy sector, Markov-chain, Kernel estimation, coefficient of variation.

JEL: O47, R12, Q18

1. Introduction

Economic cohesion has been a core goal of European integration as stated in the second article of the Maastricht Treaty on European Union (European Union, 1992). On the other hand, agriculture has been crucial for European integration since the foundation of the EEC in 1957. Agriculture was dealt with in detail in the EEC-treaty and became a frontrunner in economic European integration (Oskam et al., 2011: 37).

In 2004 and 2007, the EU welcomed in total twelve New Member States (NMS), the so-called Eastern Enlargement. They added their agricultural sectors to the EU common market. Dairy farms of the NMS faced a productivity and competitiveness gap when they entered the common market, that is a structural farm structure disparity with the Old Member States (OMS) which had been member before 2004. While an average German dairy farmer in Sachsen had a farm net income of €87.771, an average Polish dairy farmer in Mazowsze i Podlasie had a farm net income of €7.513 (FADN, 2017). Farms in the OMS had experienced pronounced structural change characterized by exits and substantial farm size growth in the decades before (Van Berkum and Helming, 2006). In contrast, in most NMS such as Bulgaria, Poland or Romania many small farms operating only a few hectares existed during accession.

Dairy products account for the largest share in total agricultural production in the EU (i.e. 13.9%), meaning that it is the largest agricultural market of the EU (European Commission, 2012). Since the Eastern Enlargement the CAP has undergone multiple reforms that resulted in more market-orientation. Especially EU dairy policy changed with the quota abolition in 2015 as well as the EU dairy sector (Jongeneel and Van Berkum, 2015).

Salou et al. (2017), Huettel and Jongeneel (2011) or Tonini and Jongeneel (2009) examine the effects of structural change of the EU dairy sector. To the best of our knowledge, whether structural convergence of regional dairy sectors has not been addressed so far. Jambor et al. (2016), Alexiadis (2010), Rezitis (2010) or Brasili et al. (2006) examine convergence between the entire agricultural sectors of the EU countries (but they did not assess single sub-sectors). Only Cechura et al. (2017) and Jansik and Irz (2015) analyse convergence of productivity of dairy production across Member States.

The Common Agricultural Policy (CAP) had an important effect on convergence and cohesion between the EU regions as stated the CAP objectives in European Union (2012). The absence of cohesion could hamper the 'continuation of the integration process' (Kuokannen and Vihinen, 2006: 6). Van Berkum (2009) and others expected the farm sector in the NMS to experience a similar structural change eliminating the substantial competitiveness gap existing during the Eastern Enlargement. To what extent did this unified EU policy scheme actually

reach this goal of cohesion in agricultural markets? Given the importance of the dairy sector in EU agriculture, we analyse to what extent the regional dairy sectors of the NMS and the OMS experienced a structural convergence since the Eastern Enlargement. Did milk yields and farm income in NMS such as Bulgaria become more similar to OMS levels such as Denmark's one or did the productivity gap endure or even grow? Our contribution consists in performing a multidimensional assessment of to what extent national raw milk prices, regional average dairy farm productivity and income in the NMS experienced a catch-up during the last decade.
In the next section the methodology and empirical models are explained, which is followed by a section on the data that is used, subsequently the results are presented. In the last sections the conclusion, implications and limitations are set out.

2. Methodology

Dunford and Smith (2000:173) state that: *"Cohesion depends on the degree of equality in the distribution of GDP per head and the extent to which there are processes of catch-up in which less developed countries and region and lower-income groups enjoy faster rates of income growth than more developed areas or richer groups."* Convergence can therefore be defined as 'increased cohesion', that is, reduced dispersion or dissimilarity as illustrated in Figure 1.

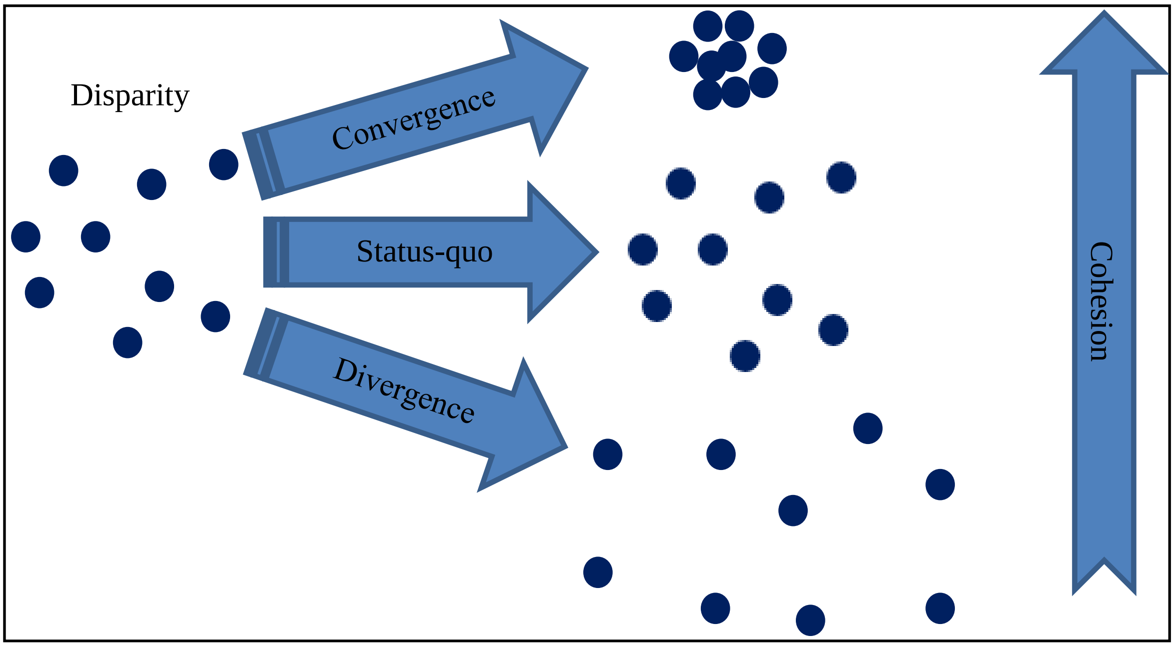

Figure 1: Disparity, convergence and cohesion illustrated
(Source: Authors)

Convergence has for a long time attracted the interest of economists. Whether poorer regions or nations could come closer to the richer ones is a crucial question for society. Most convergence studies are covered within the so-called Neo-classical Growth Theory. The first economic model that deals with convergence was developed by Robert Solow (1956). This neoclassical model has been the workhorse of economists to study convergence. The main equation of this model is:

$$Y = F(K, L) \qquad (1)$$

In which the output of an economy, Y is a function of capital, K and labour, L. One assumption of the model was that an economy diminishing returns to capital, which eventually led to the β −convergence hypothesis. Countries that were closer to the steady state had less economic growth then countries that were further away from the steady state level. In the nineties several scholars criticized the β-convergence definition for *Galton's fallacy*[1] and they proposed the use of the σ-convergence definition (Friedman, 1992; Quah, 1993)[2]. σ-convergence is a measure of the standard deviation or more specifically: the degree of dispersion across income levels or income growth rates. If the degree of dispersion reduces over time, we can speak of σ-convergence. In this study we follow the σ-convergence, so we look how the cross-section distribution of the regional dairy sectors has changed over time.

Price convergence

Many studies have examined price convergence, though the theoretical framework and empirical approaches that were applied could be rather different. The absence of a strong single theoretical framework for price convergence has resulted in multiple empirical measures of convergence as can be seen in Table 1.

Table 1: Empirical methods of price convergence[3]

Measure	Visual/Quantitative	Main characteristics
Unit root analysis	Quantitative	Convergence in the short-run dynamics. Uses the panel-data aspect in price series.
Coefficient of variation	Quantitative /Visual	Unit-free measure, based on a cross-section of data in a specific year. Regression tests can be used to view the long-run development of this variation measure.
Philips-Sul method	Quantitative	Novel approach in convergence tests. Uses the panel-data aspect in price series. Able to identify convergence clusters.
Rogers F-test	Quantitative	Based on an F-test of two moments in time.

(Source: Authors)

[1]Also known as reversion to the mean. Explained in more detail in: Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics:* 427-443.
[2] The wide spread of theoretical and empirical approaches to convergence have also led to plurality in the definition of convergence. Islam (2003) provides an excellent overview of the scientific debate on economic convergence.
[3] For more background on this emprical approaches, please see the papers of: Rogers (2007), Philips and Sul (2007) and Goldberg and Verboven (2005).

In this paper we use the coefficient of variation as a measure for price convergence, because it is a straightforward measure which can show the long-run development of price dispersion, which is considered adequate for our purpose. Unit root analysis is not suitable since this measure makes use of the short-term price dynamics to test for convergence, while our particular interest lies in the long-run. The Philips-Sul method has the disadvantage that empirical applications are limited, which makes it hard to cross-check results. The Rogers F-test is a too sensitive measure since it only uses two points in time.

The CV is a measure used to estimate price dispersion is defined as the standard deviation divided by the mean at time t (Monfort, 2008):

$$CV_t = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{i,t} - \overline{P}_t)}}{\overline{P}_t} \tag{2}$$

With $P_{i,t}$, as the price in euros of country i in month t and $\overline{P}_t$ is the average price of all n countries. As the individual prices come closer to the mean price, the dispersion decreases, hence σ-convergence occurs.

To measure whether the dispersion has decreased over time we use an OLS-regression to test for σ-convergence as defined in Gil-Pareja and Sosvilla-Rivero (2008). This is a simple regression of CV_t verses a linear time trend t with a constant α and the error term ε_t:

$$CV_t = \alpha + \sigma t + \varepsilon_t \tag{3}$$

In the past decade there existed severe periods of price volatility, which could interfere in the continuous process of convergence. Therefore, a structural break test is applied, to test whether there is a structural break present. It is found that the null-hypothesis of no structural break was rejected at 1%-level for every group of countries. For that reason, we include a dummy variable (D) and an interaction variable $(D \cdot t)$ to account for the structural break[4]:

$$CV_t = \alpha + \sigma t + \gamma D + \beta(D \cdot t) + \varepsilon_t \tag{4}$$

[4] Additionally, we have applied a log-linear and linear-log functional form. Above that, we have estimated a model with a dummy based on actual SMP intervention in the market, but these approaches were not more consistent than the linear model. Output of these models are available at request at the corresponding author.

Productivity and income convergence

Monfort (2008) provides an excellent overview of the measures that are available to assess convergence (Table 2). These measures, which are in particular used for assessing convergence of GDP per capita across regions, can equally be applied to measure convergence of productivity and farm income.

Table 2: Measures of inequality and convergence

	Measure	**Visual/Quantitative**	**Main characteristics**
Beta-convergence	Beta-coefficient	Quantitative	Estimated rather than computed
Sigma-convergence	Coefficient of variation (CV)	Quantitative	Sensitive to changes in the mean, in particular when the mean value is near zero
	Gini index	Quantitative	Sensitive to changes in inequality around the median/mode
	Atkinson index	Quantitative	Weight given to gaps between incomes in lower or upper tail of the distribution parameterised through the "aversion to inequality".
	Theil index	Quantitative	Gives equal weights across the distribution.
	Mean Logarithmic Deviation	Quantitative	Gives more weight to gaps between incomes in the lower tail of the distribution.
Analysis of distribution	Salter graphs	Visual	No possibility of statistical inference. Possibility to identify individual regions.
	Kernel estimation	Visual	No possibility of statistical inference. No possibility to identify individual regions.
	Markov chain analysis	Quantitative	Possibility of statistical inference and of identifying individual regions.
	Cumulative frequency	Visual	No possibility of statistical inference. No possibility to identify individual regions

(Source: Authors, based on Monfort, 2008: 20)

The particular interest of productivity (income) convergence is to find whether regions with low productivity (income) are catching up with the regions with the highest productivity (income). The summary statistics of σ- and β-convergence measures cannot identify dynamics within the distribution of the sample. For that reason we focus on the analysis of distribution, since this it gives a useful insight in the dynamics of the convergence process. Kernel density plot estimation is a useful tool to see whether bimodality exists in the distribution. This method could show whether there exist several ‘clubs’ or ‘groups’ of regions that have the same productivity and contribute to identify the dynamics of the *external distribution.* To also statistically assess convergence, Markov chain analysis is applied as this method allows to detect the dynamics of the *internal* distribution.

Kernel density estimation is a non-parametric technique to show the density of a distribution. Like a histogram it can show how a variable is distributed, but in a smooth way without sub-intervals (Monfort, 2008). The Kernel density estimator of a series X with a specific point x can be defined as (Silverman, 1986: 4):

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right) \tag{5}$$

n is the number of observations, k the kernel function, and h is the smoothing parameter. Like Fingleton and López-Bazo (2003), Quah (1997), Hansen and Teuber (2011) the Gaussian Kernel function (K) is used in the estimation. The optimal bandwidth h is based on the paper Silverman (1986). If σ-convergence would occur, we would expect that the distribution becomes more dense over time.

The Markov model shows a system of several 'states' in which it is possible to move from one state to the other state over time. An important property of the Markov chain is that it has no memory, i.e. that the future steps in the system from the current do not depend on the past.

The transition probability matrix **P**, shows what the probabilities are to move from one state to the other state (e.g. p_{11}, is the probability that being at state one at time t, you will stay in state 1 at time $t+1$. p_{13}, is the probability that you will move from state 1 to state 3, and p_{31} is the probability that from state 3 you will move to state 1).

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{6}$$

Matrix **P** is our object of study, because it indicates what the probability is that a region that falls within a certain productivity (income) class, is able to move up our down certain productivity (income) classes. The evolution of the distribution can be described by the transition matrix **P**:

$$\pi^{t+1} = \mathbf{P} \cdot \pi^{t} \tag{7}$$

To estimate these transition probabilities, data is required that allows us to derive these transition probabilities p_{ij}. Generally, two types of data are distinguished: micro- and macro-data. Micro-data is data that can show us the movements of the entities from one state to another over time (Zimmermann et al., 2009). Macro-data can only show the number of entities in each state at time t, so the individual movements of the entities between the states cannot be observed. In this case, micro-data is available, so we limit ourselves to the estimation procedure of micro-data. Anderson and Goodman (1957) have shown a Maximum Likelihood Estimation (MLE) procedure to estimate the transition probabilities[5]. We define $n_{ij}(t)$ as the number of individuals in state i at $t-1$ and j at time t ($i,j = 1,\dots,m$; $t = 1,\dots,T$). $n_{i(0),i(1),\dots,i(t)}$ is the number of individuals or regions for which the sequence of states is $i(0), i(1), \dots, i(T)$. We assume the transition probabilities p_{ij} to be stationary (or called homogenous over time), meaning that they do not change over time. To estimate the transition probabilities, we have to imply two restrictions (Anderson and Goodman, 1957: 92):

$$p_{ij} \geq 0 \text{ and } \sum_{j=1}^{m} p_{ij} = 1 \qquad i = 1,2,\dots,m \tag{8}$$

These restrictions imply that transition must be non-negative, hence only positive probabilities exist. The second restriction says that the sum of each probability from state $j = 1$ to the states $1,\dots,m$ should add up to one. This makes sense, since the number of movements can never exceed the number of individuals in the system. Then the Maximum Likelihood estimator for p_{ij} can be defined as (Anderson and Goodman, 1957: 92):

$$\hat{p}_{ij} = n_{ij}/n_i^* = \sum_{t=1}^{T} n_{ij}(t) \Big/ \sum_{k=1}^{m} \sum_{t=1}^{T} n_{ik}(t) = \sum_{t=1}^{T} n_{ij}(t) \Big/ \sum_{t=0}^{T-1} n_i(t) \tag{9}$$

This holds for the i-th sample ($i = 1,2,\dots,m$) that consists of $n_i^* = \Sigma_j n_{ij}$ multinomial trials with the probabilities $p_{ij}(i,j = 1,2,\dots,m)$.

[5] For more explanation regarding the statistical inference of Markov chains see: Anderson, T. W., and Goodman, L. A. (1957).

If the transition matrix **P** is stationary (or called homogenous over time) it means that all transition probabilities are equal over time (Anderson and Goodman, 1957: 92):

$$p_{ij}(t) = p_{ij} \quad \text{for all } t \tag{10}$$

From the transition matrix, several statistics on convergence can be conducted. The half-life is the number of periods it takes to close half of the gap towards the stationary distribution. The half-life can be calculated as (Shorrocks, 1978: 1021):

$$\text{Half-life} = \frac{-\log 2}{\log|\lambda_2|} \tag{11}$$

λ_2 is the second-to-largest eigenvalue of the transition matrix **P**. The half-life is value between zero and infinity. If it is zero, it means that the stationary distribution has already been reached. A second measure is the mobility index (M^{OV}), which is an indicator of the degree of mobility in the distribution. So it is a measure that indicates what the overall likelihood of the transition matrix is to remain in a certain state. If there would be no mobility in the distribution, it would mean that all probabilities along the diagonal are equal to one, hence it would be the identity matrix of **P**. Then the mobility index is equal to zero. For the case of perfect mobility, we assume a quasi-maximal diagonal for **P** (there exists a positive $\mu_1, \dots, \mu_n$ $such\ that$ $\mu_j p_{jj} \geq \mu_i p_{ij}$ for all i, j) (Shorrocks, 1978: 1017). This means that the probability to remain in the same class is not less than the probability to move to another class. So in case of perfect mobility, the trace of the probability matrix is then equal to one, consequently the mobility index will be one. This measure proposed by Shorrocks (1978: 1017) is defined in Equation 12. Where $tr(\mathbf{P})$ is the trace of the transition matrix **P** and n is the number of classes:

$$M^{OV} = [n - tr(\mathbf{P})] \cdot [n-1]^{-1} \tag{12}$$

Additionally, two mobility measures are calculated to interpret the direction of the mobility: upward or downward mobility. The sum of the upper triangle of transition probabilities represents the share of upward mobility ($\sum_i \sum_{j>i} p_{ij}$), while the sum of the lower triangle of transition probabilities represents the sum the downward mobility ($\sum_i \sum_{i>j} p_{ij}$). The mobility of the diagonal element is defined as $\sum_j \left(1 - p_{jj}\right)$. The upward and downward mobility are 'deflated' by the sum of this diagonal element and therefore $M^U + M^D = 1$. If there is no downward mobility and no persistence along the diagonal, it would mean that there is perfect upward mobility and M^U would be equal to one. If there is no upward mobility and no persistence along the diagonal, it would mean that there is perfect downward mobility and M^D would be equal to one. The upward mobility can therefore be defined as (Huettel and Jongeneel, 2011: 513):

$$M^U = \left[\sum_i \sum_{j>i} p_{ij}\right] \cdot \left[\sum_j \left(1 - p_{jj}\right)\right]^{-1} \tag{13}$$

And the downward mobility is defined as (Huettel and Jongeneel, 2011: 513):

$$M^D = \left[\sum_i \sum_{j<i} p_{ij}\right] \cdot \left[\sum_j \left(1 - p_{jj}\right)\right]^{-1} \tag{14}$$

If a process of convergence would happen, it is expected that the probabilities to move from the lowest category to higher categories in the transition matrix **P** are substantial. At the same time, the transition probabilities to move from middle categories to lower categories should not be high. If the probabilities to stay at the tails of the distribution are considerably higher than the other probabilities along the diagonal, there is a high risk that the distribution moves towards a twin-peak distribution, so that a persistent structural gap between two groups will exist.

For this study we divide the FADN regions into specific productivity (income) classes. To identify these classes, we average the productivity numbers of the regions by the sample-average. This has two advantages: fluctuations in the whole dairy market are eliminated (like EU crisis years) and it makes it easier to compare regions with each other. The choice of the number and width of the classes is rather sensitive. The higher the number of classes, the better the density of the distribution is approximated, but it comes with less reliable transition probabilities. The lower the number of classes, the rougher the division of the distribution is and the less information is abstracted from the distribution (Geppert and Stephan, 2008). We chose to divide the sample in five classes, with equal steps like Monfort (2008) and Pellegrini (2002) did.[6]

[6] *STATA 14.1* is used to estimate Kernel density plots. The R-package *markovchain* is used to estimate the transition probability matrices.[6]

3. Data

Farm –gate milk prices data

We start out with an analysis of the monthly raw milk price in € per 100kg, as provided by the Milk Market Observatory (MMO) (2017). Price data is available only at a national level, so in this part our focus is on convergence between the MS. In this analysis we use only the time series after 1996, because since that year all OMS had price data available. The NMS that accessed in 2004, had price data since 2003, Hungary and Czech Republic even since 2002. Bulgaria has data from 2007 onwards, Romania from 2009 onwards. Since we divide the group in several groups of MS, we exclude Bulgaria (2007, 2008), Hungary (2002) and Czech Republic (2002) from our analysis. This is a loss of information but it improves the consistency of the analysis. Since Malta had only data from 2011 onwards, and Croatia from 2013 onwards, we exclude these two countries from the analysis.

First, we calculate the CV for three groups (OMS, NMS (2004), NMS (2004+2007)) and visually inspect the CV over time. Subsequently, for four specific groups the CV is calculated and a convergence test is applied to these groups. It is chosen to make a distinction between large and small producing countries to see whether there is a difference in the convergence process. It could be that the largest producing countries have more complete mechanisms to transmit price shocks. The selection between large and small producing countries is based on the total collection of cow's milk per MS in 2005 (Eurostat, 2017). Countries with a higher than median collection of cow's milk are considered as large, the others small. Remarkably, the 12 largest producing countries consist of 11 OMS and Poland as the only NMS. This is in line with the findings of Ihle et al. (2017: 61), who find that the OMS account for 86% of the total milk delivered to dairies in the EU.

Productivity data

Following Jansik and Irz (2015) two measures for labour productivity are applied: output per dairy cow and cows per worker. The first one can reflect innovations in a biological sense, like breeding and genetics or feed input to improve milk yield. The second one can show underlying growth in the mechanical sense, like milking machines, feed robots or tractors. Together they form an identity for labour productivity:

$$\frac{Yield}{Labour} = \frac{Yield}{Cows} \cdot \frac{Cows}{Labour} \quad (15)$$

$$Biological\ productivity = \frac{Milk\ yield\ (kg)}{Number\ of\ cows} \tag{16}$$

$$Mechanical\ productivity = \frac{Number\ of\ cows}{Labour\ input} \tag{17}$$

The FADN database provides several input and output variables from which partial productivity indices can be constructed. The dataset that is used is a dataset from 2004-2015, with year, country, region and TF14 classification. So with this database we have aggregated data on regional FADN-level that allows to only select the specialist dairy farms (TF:45). For the biological productivity measure, we use the FADN variable: SE125 Milk Yield. It is defined as *the average production of milk and milk products (in milk equivalents) per dairy cow*. For the mechanical productivity measure, we divide: SE085 Dairy cows by SE010 Total labour input. SE085 *includes female bovine animals (including female buffaloes) which have calved and are primarily held for milk production for human consumption, cull dairy cows excluded*. SE010 *is defined as total labour input expressed in Annual Working Unit (AWU) (AWU=full-time person equivalent)*. (FADN, 2018b)
In total we include all regions that appear consecutively in the sample. This ensures that the mean value is not influenced by regions that appear only for one or two years. If for example, in 2008 a region shows up hat has a high productivity value, it could move the mean value for the group up, which then leads to movements between classes which are only the result of the region being in the sample. Since Bulgaria and Romania joined in 2007, these countries also only have data from 2007 onwards. Therefore the empirical model is conducted for two time periods. First for all the regions that have data from 2004 onwards. Secondly, for all the regions that have data from 2007 onwards. It was chosen to not take into account OMS regions that have consecutive data from for example 2006 onwards. This to make sure that we can clearly see what the effect on convergence is if we add the NMS(2007) regions. The list of FADN-regions included in the sample can be found in Appendix I.

Farm income data

Given the heterogeneity of the EU dairy sector a measure is needed that allows to compare industrial-based farming systems that uses a lot of labour with small-scale self-sufficient farming systems.

For that reason, the first income measure applied is the *labour-adjusted value added*:

$$Labour - adjusted\ value\ added = \frac{Farm\ net\ value\ added}{Labour\ input} \qquad (18)$$

In the FADN database the variable SE425 Farm Net Value Added/AWU can be used as the variable for Labour-adjusted net value added. It is defined *as the Farm Net Value Added per Annual Working Unit*.

As second income indicator we use the Farm net income instead of the Farm net value added. The farm net value added is corrected for external factor costs and for subsidies and taxes on investments. We define this variable as *labour income*:

$$Labour\ income = \frac{Farm\ net\ income}{Labour\ input} \qquad (19)$$

In the FADN database the variable can be calculated by dividing SE420 Farm Net Income by SE010 AWU (Annual Working Unit).

As third income indicator we take a closer look at the family aspect in farming. In many regions in the EU farming is still organized as a family business. Usually (unpaid) family labour is a major component in the dairy farm. As a third income measure we therefore propose *family farm income*:

$$Family\ farm\ income = \frac{Farm\ net\ income}{Family\ labour} \qquad (20)$$

Family farm income can provide insight into the family component present in dairy farming. It is important since a lot of families depend on their farm for their income, so it can measure an important living condition for the families. In the FADN database it is defined as: SE430 Family Farm Income / FWU. The variable is described as *family farm income expressed per family labour unit. Takes into account difference in the family labour force to be remunerated per holding. It is calculated only for the farms with family labour.* (FADN, 2018b).

In total 84 FADN regions are in the sample for 2004-2015 and 91 FADN regions are in the sample for 2007-2015.

4. Results

Price convergence

In Figure 2, the CV for the three different MS sub-groups can be found. In this graph our interesting patterns in the dispersion are identified The first pattern is that between 2003 and 2007, the prices of the 9 NMS (Malta excl.) that entered in 2004, came closer to the OMS. At the end of 2007, the difference between the dispersion of OMS and the dispersion of OMS+NMS(2004) is very small (grey area). Secondly, we can see that the inclusion of Bulgaria and Romania did not change the dispersion significantly. Third, it can be seen that when there is a price decline in the EU milk market then the dispersion increases (marked yellow). In addition, the gap between the dispersion of OMS and the dispersion of OMS+NMS grows larger when there is a price decline in the market. The last pattern that can be observed is that up to 2008 price dispersion is quite stable, except for some seasonal variation. After 2008, the dispersion fluctuates also more, probably to non-symmetrical price changes in the national milk prices.

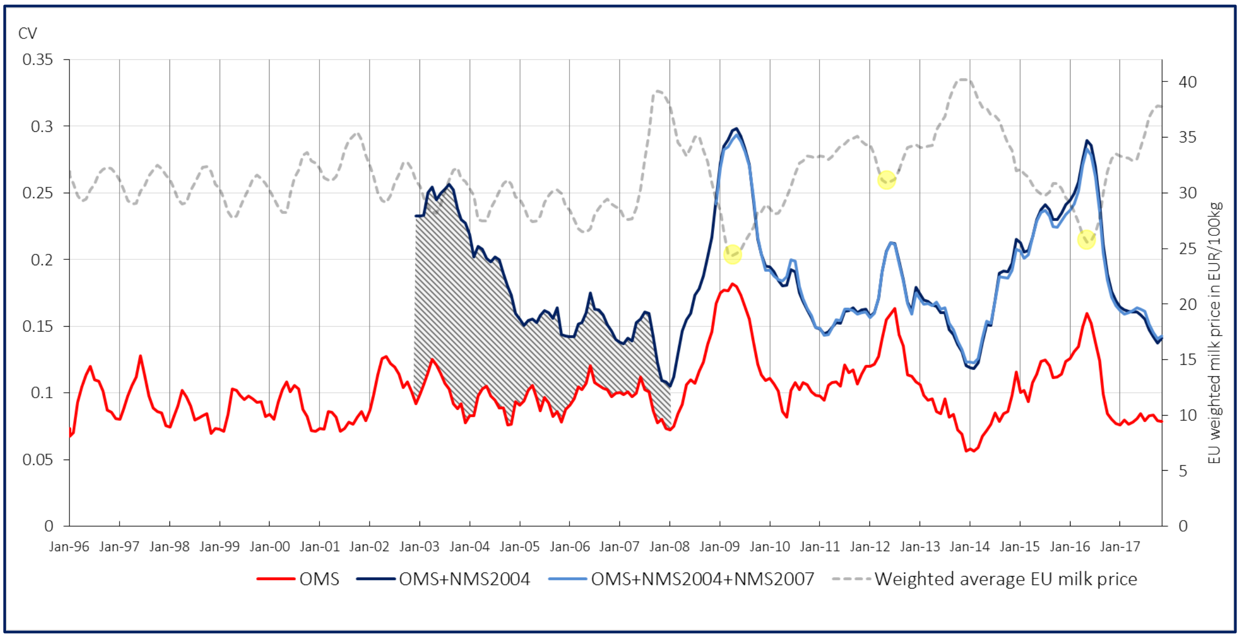

Figure 2: Coefficient of variation for 3 MS subgroups between 1996 and 2017
Note: The CV is represented on the left vertical axis. On the right vertical axis the weighted EU raw milk price can be found, this is measured in €/100kg
(Source: Authors, based on data from MMO, 2017)

In Figure 3, the CV for the groups of large- and small- dairy producing countries are shown. Remarkable is that the CV of the largest producing countries is quite stable over time. The CV of the largest producing countries is even lower than the CV of the OMS. The prices of the largest producing countries are moving together, resulting in a more or less stable CV. An explanation for this phenomenon could be that there is considerable price transmission between these large markets. Since these countries are major players on the EU- and world market, the price transmission is more complete than for the smaller producing countries. In 2014, when the Russian trade ban came into place, price dispersion amongst the smallest producers increased substantially, while the largest producing countries only show a minor increase in the dispersion.

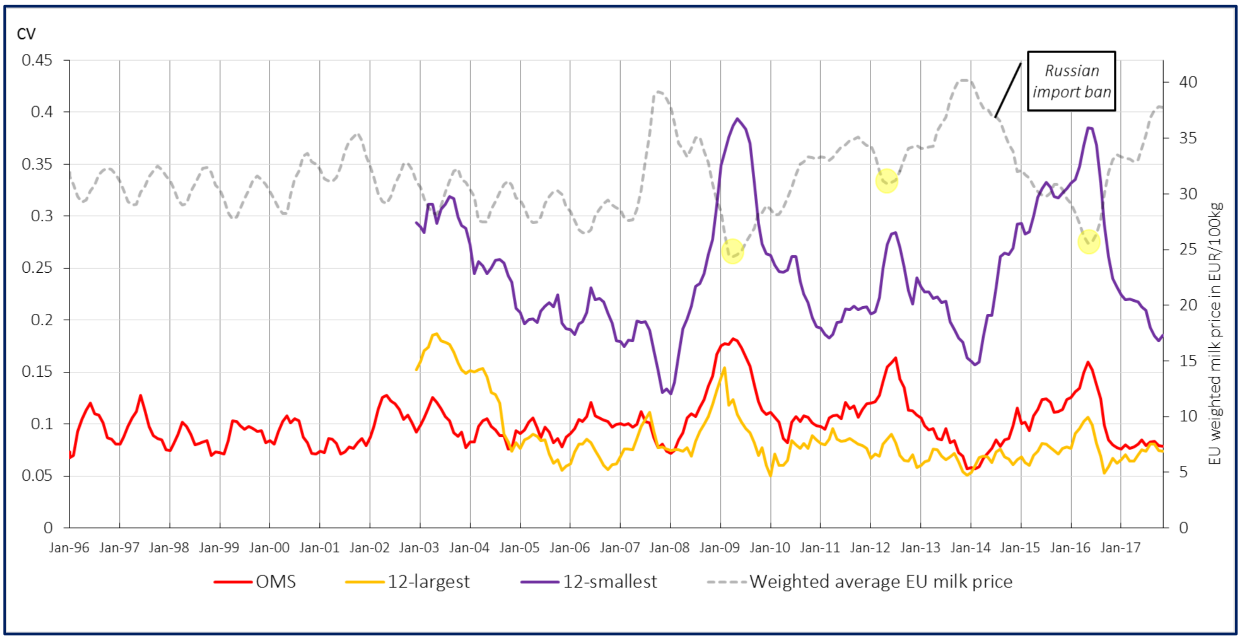

Figure 3: Coefficient of variation for 3 subgroups between 1996 and 2017
Note: The CV is represented on the left vertical axis. On the right vertical axis, the weighted EU raw milk price can be found, this is measured in €/100kg.
(Source: Authors, based on data from MMO, 2017)

In Table 3, the results of the linear regression with one dummy variable for the structural break are shown. For all four groups the variable t is significant at 1%-level. From one-sided t-tests, we find that the null-hypothesis of $\sigma \geq 0$ was rejected at 1%-level. So, we can expect this coefficient to be negative for all groups of countries, meaning that *ceteris paribus* the CV has decreased over time, hence dispersion decreased, which means that price convergence has occurred. The group of the 12 largest countries have the lowest coefficient for t, which would mean that this group has faced the strongest price convergence process. The interaction variable $t \cdot D$ is significantly different from zero for 3 out of 4 models at 1%-level. This would mean that indeed the time effect on the CV is different after the structural break. For the OMS+NMS(2004&2007), the interaction variable was not significant. This can be explained by the fact that this variable only has values from 2009 onwards. The dummy variable is only significantly different from zero for the 12 largest countries, which indicates at a clear level shift after the structural break. They show reasonable values, with the lowest value for the 12 largest countries (i.e. 0.19333), in Figure 3 it was also seen that this variable shows the lowest CV in general. The model statistics show clearly that the model for the 12 largest countries seems to fit the best, which is reflected in the highest F- and R^2 values and the lowest AIC value. The results should be treated with the utmost caution given the fact that inspection of the residuals and several tests showed that the regression seems to fail the Gauss-Markov assumptions.

Table 3: Regression results of the linear model

Variables	$OMS + NMS(2004)$	$OMS + NMS(2004\&2007)$	$12\ largest$	$12\ smallest$
t	-0.00192***	-0.00157***	-0.00364***	-0.00222***
	(0.00)	(0.00)	(0.00)	(0.00)
D	-0.02837*	0.12888***	-0.10247***	-0.02542
	(0.02)	(0.04)	(0.01)	(0.02)
$t \cdot D$	0.00177***	-0.00026	0.00353***	0.00208***
	(0.00)	(0.00)	(0.00)	(0.00)
$constant$	0.23885***	0.23752***	0.19333***	0.29856***
	(0.01)	(0.01)	(0.01)	(0.01)
n	*180*	*108*	*180*	*180*
F	*24.12875*	*29.56878*	*156.56350*	*22.17043*
R^2	*0.29143*	*0.46032*	*0.72742*	*0.27426*
AIC	*-664.63791*	*-428.37694*	*-968.44545*	*-557.49285*

Note: estimates are rounded to 5 decimals, *** significant at 1%-level, ** significant at 5%-level *significant at 10%-level (Source: Authors, based on data from MMO, 2017)

Productivity convergence

As can be seen in Figure 4, there is a considerable variation in the biological productivity. During the period 2007-2015, bimodality appears in the distribution. In 2015, there is a clear peak around 3000 kg, and a second peak at 7500 kg. A second process is observable in the second peak, which is slightly shifting to the right over time.

For mechanical productivity, again a bimodal or twin-peak distribution can be observed. For the years 2004-2006 this bimodality is already present. In the period 2007-2015 the sharp peaks disappear, resulting in a wider spread. The largest peak has shifted to the right over time, while the smallest peak stays at the same position. The spread of the distribution has not substantially changed over time. Since the twin-peak distribution stays in place, there are no signs of convergence present in the distributional dynamics of the Kernel density plot.

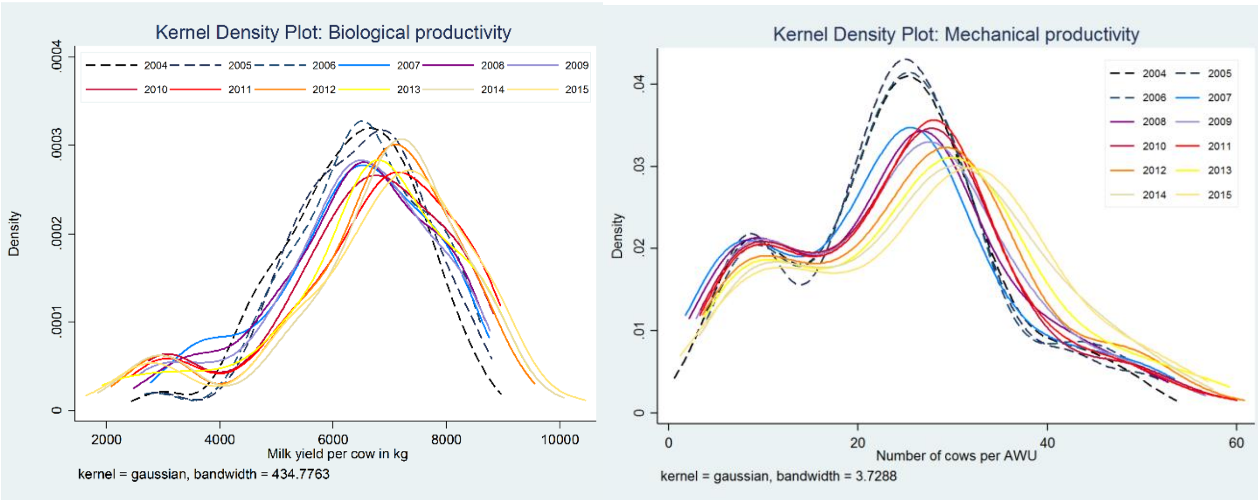

Figure 4: Kernel density plots of biological and mechanical productivity between 2004-2015
(Source: Authors, based on data FADN, 2018a)

Table 4 provides the estimates of the Markov transition probability matrix. First it can be noted that the transition probability to remain in the lowest class is high for all variables and time periods. Secondly, the results show the probability to move from the lowest to the one-to-lowest category is low (i.e. in 3 out of 4 cases <10%). The probability to move from the one-two-lowest category to the middle category is somewhat higher (i.e. 10-15%). If we add the regions that accessed in 2007, the transition probabilities change considerably. The mobility index is low for all cases, since it is closer to zero than to one. Next to that, we find that upward mobility is higher than downward mobility in all cases. Comparing the two variables, it can be found that biological productivity has a lower half-life and higher mobility index. This means that the process towards convergence is faster for biological productivity and there is less persistence in the distribution. Comparing the two time periods, the results show that there is more persistence in the 2007-2015 period, compared to the 2004-2015 period. Adding the 2007 NMS regions, results in higher transition probabilities for the lowest class, a higher half-life and a lower mobility index, meaning that convergence in this period is less and slower.

Table 4: Statistics of the transition probability matrices of biological and mechanical productivity for FADN sample regions from 2004-2015

Variable	Biological productivity		Mechanical productivity	
	2004-2015	*2007-2015*	*2004-2015*	*2007-2015*
p_{11}	0.833	0.929	0.925	0.956
p_{12}	0.154	0.071	0.075	0.044
p_{23}	0.140	0.103	0.142	0.103
Half-life	7.877	15.882	13.084	19.866
Mobility index	0.234	0.191	0.168	0.154
Upward mobility	0.552	0.547	0.565	0.553
Downward mobility	0.448	0.453	0.435	0.447

Note: estimates are rounded to three decimals. Detailed results can be found in Appendix II.
(Source: Authors, based on data FADN, 2018a)

Income convergence

As can be seen in Figure 5, the Kernel density plots vary much more by year than in the case of productivity. Still, for most years a pattern can be observed. For labour-adjusted value added, there is a small peak at the beginning of the distribution followed by a very large peak. The distribution is skewed to the right, perhaps due to the case that income is bounded to zero. For the year 2009, the distribution is completely different, since the distribution is shifted to the left. This corresponds to the dairy crisis in that year which implied a shift to lower levels of value added. Like in the productivity case we see a shift of the larger peak to the right (see dashed arrow), while the peak at the left stays at the same point.

For labour income, there is a large difference between the density plots of the years 2004-2006 and 2017-2015, because the distribution is less dense at the middle. Instead, from 2007 onwards the spread of labour income is wider. In the years 2011-2015, the first small peak at the lefts seems to have disappeared, which could hint towards convergence. Like in the plot of labour-adjusted value added, it is skewed to the right.

The Kernel density plot for family farm income suggests that there is not a stable distribution over time. Moreover, bimodality cannot be seen in this distribution. The main peak has moved slightly to the right, which means that for a large group the family farm income has grown. The figure also shows that negative family incomes are present in the sample. Except for the year

2009, the short left tail suggests that there are only a few regions that face a small negative family farm income.

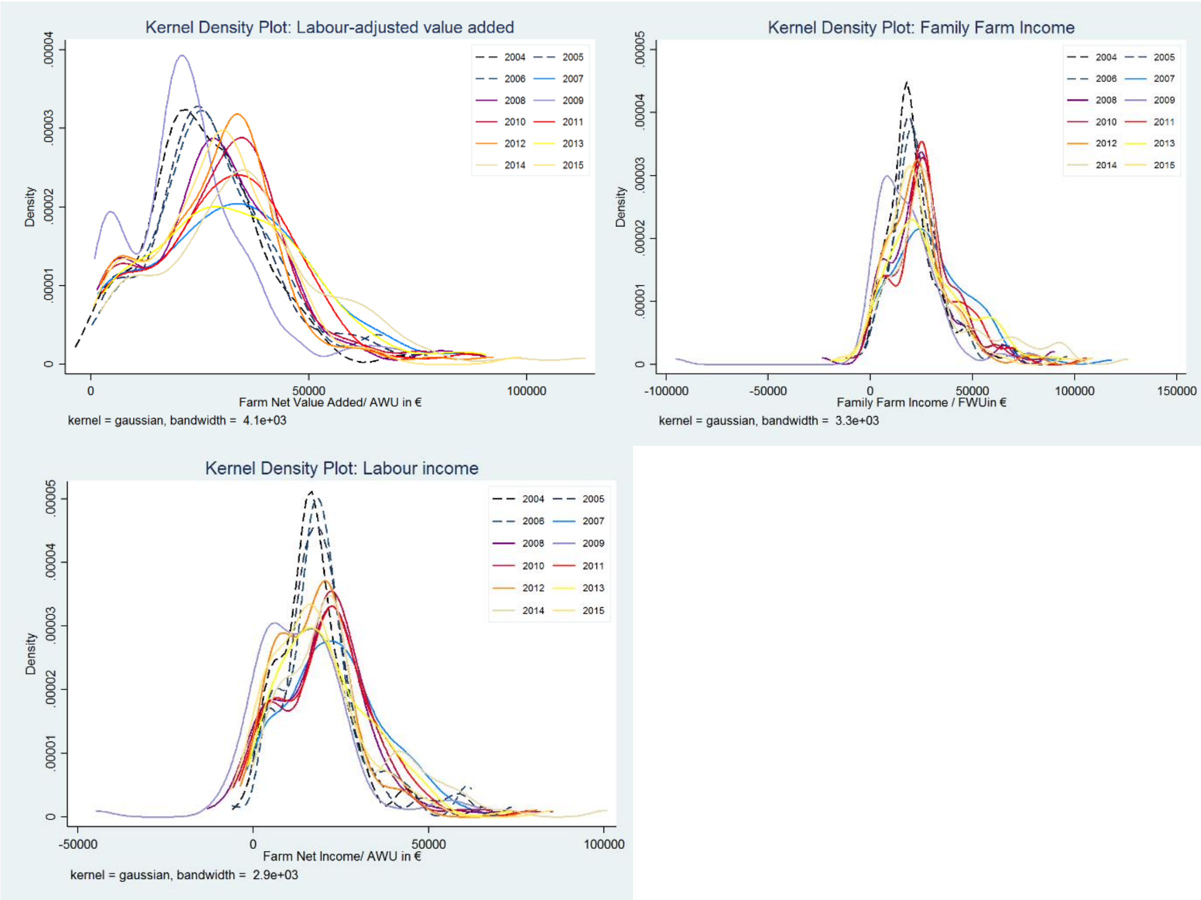

Figure 5: Kernel density plots of labour-adjusted value added, labour income and family farm income between 2004-2015
(Source: Authors based, on data FADN, 2018a)

In Table 5 the transition matrix of the labour-adjusted added value from 2004 onwards is shown. Similar to the productivity case, the probability to stay in the first class is high, although it is lower now. The probability to move from the lowest to the one-to-lowest income class is again low (i.e. around 0.1). The probability to move from the one-to-lowest class to the middle class is considerably higher than in the case of productivity. The half-life of all variables are lower than the half-lives of productivity, particularly for family farm income (i.e. 3 years). Compared to the transition matrices of the productivity variables, there is a lower half-life and higher mobility index. The lower half-life implicates that it takes less time to convergence to the stationary distribution. A higher mobility index indicates that the process is less stable which implies that the probability to move from one category to another category is higher overall. As was found in the productivity transition matrices, the probability to stay in the first class is higher for the period 2007-2015 compared to 2004-2015. Above that, p_{12} is lower for the 2007-2015 period and the half-life is also higher. Opposed to productivity, the mobility index is higher for 2007-2015, meaning that there was a higher overall probability to move from one category to another. This could be due to the fact that after 2007, price volatility entered the market.

Table 5: Statistics of the transition probability matrices of labour-adjusted value added, labour income and family farm income for FADN sample regions from 2004-2015

Variable	Labour-adjusted value added		Labour income		Family farm income	
	2004-2015	*2007-2015*	*2004-2015*	*2007-2015*	*2004-2015*	*2007-2015*
p_{11}	0.889	0.918	0.816	0.876	0.711	0.812
p_{12}	0.111	0.082	0.116	0.062	0.189	0.079
p_{23}	0.232	0.313	0.336	0.337	0.281	0.341
Half-life	6.652	8.091	3.516	4.257	2.927	3.073
Mobility index	0.389	0.454	0.535	0.582	0.522	0.551
Upward mobility	0.492	0.522	0.507	0.525	0.526	0.474
Downward mobility	0.508	0.478	0.493	0.475	0.474	0.526

Note: estimates are rounded to three decimals. Detailed results can be found in Appendix II.
(Source: Authors based on data FADN, 2018a)

7. Conclusions

Economic cohesion as well as agriculture have been playing central role in the first seven decades of European integration. Therefore, we asses to what extent convergence between the structures of the dairy sectors in EU regions took place since 2004. When the NMS entered the EU, disparities in the economic structure of dairy producers have risen considerably. We find that these disparities did largely not decrease until 2017. We find strong evidence for convergence of national raw milk prices in OMS and NMS in the first four years after the Eastern Enlargement since price dispersion between the OMS and NMS decreased substantially under stable market conditions. For structural convergence for productivity and income we find barely evidence as the Markov chains estimated show high persistence and the Kernel density plots show pronounced and stable bimodality. For income convergence evidence was stronger than for productivity convergence. For the period 2007-2015 considering also Bulgarian and Romanian regions, convergence was weaker compared to 2004-2015 without these regions.

The results of our analysis which considers FADN regions and extents the time frame to the period after the Eastern enlargement is in line with previous findings of the literature. Cechura et al. (2017) found no evidence for convergence in terms of productivity for the period 2004-2011. Jansik and Irz (2015) found that total factor productivity of the NMS did not converge to the level of the OMS. Brasili et al. (2006) found that family farm income shows more convergence than net value added per ha.

Our findings imply for EU policy making that convergence is a very slow process, which is likely to be related to (partial) structural adjustments over time. EU policy measures that facilitate such adjustments are probably the best way to promote further convergence. Measures focusing exclusively on income transfers (e.g. direct payments) may hamper structural adjustment, even though short-run impacts on convergence in farm income might be positive. However, our analysis indicates that convergence of productivity has been weaker than convergence of income. This suggests that policy makers may need to make a choice. This choice consists in whether EU policies want to strengthen the viability of (dairy) farming in the NMS by improving the level and stability of income (which might reduce incentives for productivity convergence) or by improving competitiveness of (dairy) farms in the NMS.

The analysis provided in this study has a number of limitations. First, we have chosen in this thesis to focus on specialist dairy farms, which ignores the mixed farm types that still are much present in especially the NMS. Moreover, there are some weaknesses in the FADN data. Out of 148 FADN, only 91 regions are in the sample, which can result in selection bias. Some regions also have very small sample sizes and therefore they do not give a good representation of the population. Also the fact that we measure convergence across territorial units (FADN) can be criticized on the fact that not every region has an equally large dairy sector. A counterargument for this is that the policy and policy debate is often linked to territorial units. A drawback of the CV method is that we are not sure about the functional form of the model. The market volatility makes it difficult to empirically find evidence for convergence, since the volatility troubles the true convergence process. Tests and plots on the residuals show that the model is far from perfect. The Kernel density plot could suffer from outliers which can change the distribution substantially (e.g. the crisis year 2009). A particular weakness of the Markov chain is that we arbitrarily chose the several classes. Although the classes are chosen by inspection of the quantiles and shares between the classes are equal, changing the class changes the transition probabilities. Secondly, we assume stationarity in the Markov chain, however it can be questioned whether transition probabilities are identical over a turbulent period of twelve years.

References

- Alexiadis, S. (2010). Convergence in agriculture: Evidence from the European Regions. *Agricultural Economics Review,* 11(2): 84-96.
- Anderson, T. W., and Goodman, L. A. (1957). Statistical inference about Markov chains. *The Annals of Mathematical Statistics*: 89-110.
- Brasili, C., Fanfani, R., & Gutierrez, L. (2006). Convergence in the Agricultural Incomes: A Comparison between the US and EU. Contributed Paper to the International Association of Agricultural Economists Conference Gold Coast, Australia, August 12-18,2006.
- Cechura, L., Grau, A., Hockmann, H., Levkovych, I., and Kroupova, Z. (2017). Catching Up or Falling Behind in European Agriculture: The Case of Milk Production. *Journal of Agricultural Economics*, 68(1) : 206-227.
- Dunford, M., and Smith, A. (2000). Catching up or falling behind? Economic performance and regional trajectories in the "New Europe". *Economic Geography*: 76(2), 169-195.
- European Commission. (2012). Agriculture in the European Union Statistical and Economic Information. Available from: https://ec.europa.eu/agriculture/sites/agriculture/files/statistics/agricultural/2012/pdf/full-report_en.pdf [Accessed: 02-07-2017]
- European Commission. (2017). Report on the distribution of direct payments. Available from: https://ec.europa.eu/agriculture/sites/agriculture/files/cap-funding/beneficiaries/direct-aid/pdf/annex2-2016_en.pdf [Accessed: 18-02-2018]
- European Union. (1992). Treaty on European Union (Consolidated Version), Treaty of Maastricht. *Official Journal of the European Communities C 325/5; 24 December 2002*. Available from: http://www.refworld.org/docid/3ae6b39218.html [Accessed: 12-02-2018]
- European Union. (2012). Consolidated version of the Treaty on the Functioning of the European Union: Article 39. OJ C 326, 26.10.2012: 47–390.
- Eurostat. (2017). Collection of cow's milk in 1000 tonnes. Available from: http://ec.europa.eu/eurostat/tgm/download.do?tab=table&plugin=1&language=en&pcode=tag00037
- FADN. (2017). Farm Net Income (SE420)FADN Public Database: Year(2004), Country(All), Regions(All), Type of farming(5: Milk). Available from: http://ec.europa.eu/agriculture/rica/database/report_en.cfm?dwh=SO [Accessed: 02-07-2017]

- FADN. (2018a). Standard Results Database: YEAR.COUNTRY.REGION.TF14.zip. Available from: http://ec.europa.eu/agriculture/rica/database/consult_std_reports_en.cfm [Downloaded: 11-01-2018]
- FADN. (2018b). Definition of variables. Available from: http://ec.europa.eu/agriculture/rica/definitions_en.cfm [Accessed: 9-1-2018]
- Fingleton, B., and López-Bazo, E. (2003). Explaining the distribution of manufacturing productivity in the EU regions. In: Fingleton, B. (ed.). *European Regional Growth*. Berlin, Heidelberg: Springer, 375-409.
- Fischer, C. (2012). Price convergence in the EMU? Evidence from micro data. *European Economic Review*, 56(4) : 757-776.
- Friedman, M. (1992). Do Old Fallacies Ever Die? *Journal of Economic Literature,* 30(4) : 2129-2132.
- Geppert, K., and Stephan, A. (2008). Regional disparities in the European Union: Convergence and agglomeration. *Papers in Regional Science,* 87(2) : 193-217.
- Gil-Pareja, S., and Sosvilla-Rivero, S. (2008). Price convergence in the European car market. *Applied Economics,* 40(2), 241-250.
- Goldberg, P. K., and Verboven, F. (2005). Market integration and convergence to the Law of One Price: evidence from the European car market. *Journal of international Economics,* 65(1) : 49-73.
- Hansen, H., and Teuber, R. (2011). Assessing the impacts of EU's common agricultural policy on regional convergence: sub-national evidence from Germany. *Applied Economics,* 43(26) : 3755-3765.
- Huettel, S., and Jongeneel, R. (2011). How has the EU milk quota affected patterns of herd-size change? *European Review of Agricultural Economics,* 38(4): 497-527.
- Ihle, R., Dries, L., Jongeneel, R., Venus, T. and Wesseler, J. (2017). The EU Cattle Sector: Challenges and Opportunities - Milk and Meat. Research for Agri Committee, IP/B/AGRI/IC/2016-014. Available from: http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU(2017)585911
- Islam, N. (2003). What have we learnt from the convergence debate? *Journal of economic surveys*, 17(3): 309-362.
- Jambor, A., Somai, M., and Kovács, S. (2016). 10 Years of EU Membership: Diverging Performances in New Member States Agriculture. *Ekonomický časopis*, 64(5): 458-476.

- Jansik, C., and Irz, X. (2015). Competitiveness of dairy farms in northern Europe: A cross-country analysis. *Agricultural and Food Science*, 24(3): 206-218.
- Jongeneel, R. (2011). European Union dairy policy reform: impact and challenges. In: Moreddu, C. (ed.), *Disaggregated Impacts of CAP Reforms: Proceedings of an OECD Workshop*. OECD publishing, 93-110.
- Jongeneel, R., Van Berkum, S. (2015). What will happen after the EU milk quota system expires in 2015?: an assessment of the Dutch dairy sector. *LEI Wageningen UR*, (No. 2015-041).
- Kuokkanen, K., and Vihinen, H. (2006). Contribution of the CAP to the general objectives of the EU. *MTT Economic Research.*
- Milk Market Observatory. (2017). Historical EU Price Serie Of Cow's Raw Milk In Euro/100 Kg. Regulation (EU) No 2017/1185 Article 12(a) - Annex II.4(a)
- Monfort, P. (2008). Convergence of EU regions: Measures and evolution. Brussels: European Commission, Regional Policy.
- Oskam, A., Meester, G., and Silvis, H. (Eds.) (2011). *EU policy for agriculture, food and rural areas.* 2nd edition. Wageningen Academic Publishers.
- Pellegrini, G. (2002). Proximity, polarization, and local labor market performances. *Networks and Spatial Economics*, 2(2) : 151-173.
- Philips, P. C., and Sul, D. (2007). Some empirics on economic growth under heterogeneous technology. *Journal of Macroeconomics*, 29(3) : 455-469.
- Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*: 427-443.
- Quah, D. T. (1997): Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2(1), 27-59.
- Rezitis, A. N. (2010). Agricultural productivity and convergence: Europe and the United States. *Applied Economics*, 42(8): 1029-1044.
- Rogers, J. H. (2007). Monetary union, price level convergence, and inflation: How close is Europe to the USA? *Journal of Monetary economics*, 54(3): 785-796.
- Salou, T., van der Werf, H. M., Levert, F., Forslund, A., Hercule, J., and Le Mouël, C. (2017). Could EU dairy quota removal favour some dairy production systems over others? The case of French dairy production systems. *Agricultural Systems*, 153: 1-10.
- Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica: Journal of the Econometric Society*: 1013-1024.

- Silverman, B. W. (1986). *Density estimation for statistics and data analysis (Vol. 26).* CRC press.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics,* 70(1): 65-94.
- Sosvilla-Rivero, S., and Gil-Pareja, S. (2004). Price convergence in the European Union. *Applied Economics Letters*, 11(1): 39-47.
  to agricultural producers (financial year2016). Available from: https://ec.europa.eu/agriculture/sites/agriculture/files/cap-funding/beneficiaries/direct-aid/pdf/annex2-2016_en.pdf [Accessed: 16-02-2018]
- Tonini, A., and Jongeneel, R. (2009). The distribution of dairy farm size in Poland: a Markov approach based on information theory. *Applied Economics,* 41(1): 55-69.
- Van Berkum, S. (2009). Dairy chain competitiveness in EU's new Member States, candidate and potential candidate countries. *Agrarwirtschaft*, 58 (7): 314-323.
- Van Berkum, S., and Helming, J. (2006). European dairy policy in the years to come: impact of quota abolition on the dairy sector. *LEI Wageningen UR*.
- Zimmermann, A., Heckelei, T., & Domínguez, I. P. (2009). Modelling farm structural change for integrated ex-ante assessment: review of methods and determinants. *Environmental Science & Policy*, 12(5): 601-618.

Appendix I

Region name	*Code*	*Region name*	*Code*	*Region name*	*Code*	*Region name*	*Code*
Schleswig-Holstein	10	Bretagne	163	Denmark	370	Lan i norra	730
Niedersachsen	30	Poitou-Charentes	164	Ireland	380	Czech Republic	745
Nordrhein-Westfalen	50	Aquitaine	182	England-North	411	Estonia	755
Hessen	60	Midi-Pyrénées	183	England-East	412	Latvia	770
Rheinland-Pfalz	70	Rhônes-Alpes	192	England-West	413	Lithuania	775
Baden-Württemberg	80	Auvergne	193	Wales	421	Malta	780
Bayern	90	Valle d'Aoste	221	Scotland	431	Pomorze-Muzurie	785
Saarland	100	Piemonte	222	Northern Ireland	441	Wielkpolska-Slask	790
Brandenburg	112	Lombardia	230	Galicia	500	Mazowsze-Podlasie	795
Mecklenburg-Vorpommern	113	Trentino	241	Asturias	505	Malopolska-Pogórze	800
Sachsen	114	Alto-Adige	242	Cantabria	510	Slovakia	810
Sachsen-Anhalt	115	Veneto	243	Pais Vasco	515	Slovenia	820
Thueringen	116	Friuli-Venezia	244	Navarra	520	*Severozapaden*	*831*
Champagne-Ardenne	131	Emilia-Romagna	260	Baleares	540	*Severen tsentralen*	*832*
Picardie	132	Lazio	291	Castilla-León	545	*Severoiztochten*	*833*
Haute-Normandie	133	Molise	301	Andalucia	575	*Yugozapaden*	*834*
Centre	134	Campania	302	Açores e da Madeira	650	*Yuzhen tsentralen*	*835*
Basse-Normandie	135	Puglia	311	Austria	660	*Yugoiztochen*	*836*
Bourgogne	136	Basilicata	312	Etela-Suomi	670	*Nord-Est*	*840*
Nord-Pas-de-Calais	141	Sardegna	330	Sisa-Suomi	680		
Lorraine	151	Vlaanderen	341	Pohjanmaa	690		
Alsace	152	Wallonie	343	Pohjois-Suomi	700		
Franche-Comté	153	Luxembourg	350	Slattbygdslan	710		
Pays de la Loire	162	The Netherlands	360	Skogs-och mellanbygdslan	720		

Appendix II

Transition probability matrix of biological productivity (milk yield per cow) for 84 FADN regions from 2004-2015

Initial distribution (2004)	Class	*<75*	*75-90*	*90-105*	*105-120*	*>120*
8	*<75*	0.833	0.154	0.000	0.000	0.013
15	*75-90*	0.070	0.790	0.140	0.000	0.000
25	*90-105*	0.006	0.055	0.820	0.113	0.006
25	*105-120*	0.000	0.000	0.150	0.760	0.090
11	*>120*	0.007	0.000	0.007	0.124	0.862
Summary statistics						
Class		*<75*	*75-90*	*90-105*	*105-120*	*>120*
Stationary distribution		0.082	0.146	0.331	0.252	0.188
Half-life		7.877				
Mobility index		0.234				
Upward mobility		0.552				
Downward mobility		0.448				

Transition probability matrix of biological productivity (milk yield per cow) for 91 FADN regions from 2007-2015

Initial distribution (2007)	Class	*<75*	*75-90*	*90-105*	*105-120*	*>120*
13	*<75*	0.929	0.071	0.000	0.000	0.000
12	*75-90*	0.062	0.825	0.103	0.010	0.000
25	*90-105*	0.000	0.054	0.801	0.140	0.005
21	*105-120*	0.000	0.005	0.111	0.794	0.090
20	*>120*	0.000	0.000	0.006	0.108	0.885
Summary statistics						
Class		*<75*	*75-90*	*90-105*	*105-120*	*>120*
Stationary distribution		0.109	0.125	0.235	0.291	0.239
Half-life		15.882				
Mobility index		0.191				
Upward mobility		0.547				
Downward mobility		0.453				

Transition probability matrix of mechanical productivity (cows per AWU) for 84 FADN regions from 2004-2015

Initial distribution (2004)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
18	*<50*	0.925	0.075	0.000	0.000	0.000
10	*50-80*	0.039	0.819	0.142	0.000	0.000
24	*80-110*	0.004	0.040	0.839	0.113	0.004
19	*110-140*	0.000	0.009	0.127	0.816	0.047
13	*>140*	0.000	0.000	0.000	0.073	0.927
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.079				
Half-life		13.084				
Mobility index		0.168				
Upward mobility		0.565				

Transition probability matrix of mechanical productivity (cows per AWU) for 91 FADN regions from 2007-2015

Initial distribution (2007)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
22	*<50*	0.956	0.044	0.000	0.000	0.000
9	*50-80*	0.034	0.862	0.103	0.000	0.000
20	*80-110*	0.007	0.046	0.795	0.139	0.013
23	*110-140*	0.000	0.005	0.103	0.851	0.041
17	*>140*	0.000	0.000	0.007	0.073	0.920
Summary statistics						
Class		*<50*	*50-75*	*75-100*	*100-125*	*>125*
Stationary distribution		0.138	0.132	0.228	0.307	0.195
Half-life		19.866				
Mobility index		0.154				
Upward mobility		0.553				
Downward mobility		0.447				

Transition probability matrix of labour adjusted value added (farm net value added/AWU) for 84 FADN regions from 2004-2015

Initial distribution (2004)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
12	*<50*	0.889	0.111	0.000	0.000	0.000
20	*50-80*	0.070	0.654	0.232	0.038	0.005
19	*80-110*	0.008	0.151	0.660	0.154	0.027
17	*110-140*	0.000	0.021	0.234	0.548	0.197
16	*>140*	0.000	0.000	0.064	0.242	0.694
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.141	0.189	0.302	0.206	0.163
Half-life		6.652				
Mobility index		0.389				
Upward mobility		0.492				
Downward mobility		0.508				

Transition probability matrix of labour adjusted value added (farm net value added/AWU) for 91 FADN regions from 2007-2015

Initial distribution (2007)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
19	*<50*	0.918	0.082	0.000	0.000	0.000
16	*50-80*	0.061	0.591	0.313	0.026	0.009
17	*80-110*	0.012	0.182	0.506	0.259	0.041
18	*110-140*	0.000	0.033	0.285	0.464	0.219
21	*>140*	0.000	0.000	0.082	0.212	0.705
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.160	0.166	0.261	0.213	0.199
Half-life		8.091				
Mobility index		0.454				
Upward mobility		0.522				
Downward mobility		0.478				

Transition probability matrix of labour income (farm net income/AWU) for 84 FADN regions from 2004-2015

Initial distribution (2004)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
17	*<50*	0.816	0.116	0.058	0.005	0.005
11	*50-80*	0.141	0.456	0.336	0.060	0.007
28	*80-110*	0.042	0.187	0.515	0.183	0.073
12	*110-140*	0.037	0.059	0.324	0.338	0.243
16	*>140*	0.005	0.016	0.086	0.160	0.733
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.219	0.161	0.269	0.142	0.210
Half-life		3.516				
Mobility index		0.535				
Upward mobility		0.507				
Downward mobility		0.493				

Transition probability matrix family farm income (family farm income/FWU) for84 FADN regions from 2004-2015

Initial distribution (2004)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
16	*<50*	0.711	0.189	0.082	0.013	0.006
19	*50-80*	0.104	0.576	0.281	0.026	0.013
22	*80-110*	0.060	0.239	0.504	0.132	0.064
11	*110-140*	0.060	0.078	0.172	0.397	0.293
16	*>140*	0.005	0.016	0.087	0.168	0.723
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.171	0.247	0.249	0.127	0.207
Half-life		2.927				
Mobility index		0.522				
Upward mobility		0.526				
Downward mobility		0.474				

Transition probability matrix of labour income (farm net income/AWU) for 91 FADN regions from 2007-2015

Initial distribution (2007)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
19	*<50*	0.876	0.062	0.041	0.021	0.000
19	*50-80*	0.163	0.337	0.337	0.152	0.011
17	*80-110*	0.060	0.180	0.413	0.240	0.107
15	*110-140*	0.040	0.089	0.290	0.331	0.250
21	*>140*	0.006	0.024	0.101	0.154	0.716
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.307	0.113	0.202	0.158	0.220
Half-life		4.257				
Mobility index		0.582				
Upward mobility		0.525				
Downward mobility		0.475				

Transition probability matrix of family farm income (family farm income/FWU) for 91 FADN regions from 2007-2015

Initial distribution (2007)	Class	*<50*	*50-80*	*80-110*	*110-140*	*>140*
18	*<50*	0.812	0.079	0.085	0.012	0.012
21	*50-80*	0.094	0.493	0.341	0.043	0.029
20	*80-110*	0.071	0.217	0.489	0.158	0.065
10	*110-140*	0.091	0.143	0.312	0.234	0.221
22	*>140*	0.006	0.024	0.067	0.134	0.768
Summary statistics						
Class		*<50*	*50-80*	*80-110*	*110-140*	*>140*
Stationary distribution		0.247	0.187	0.256	0.103	0.207
Half-life		3.073				
Mobility index		0.551				
Upward mobility		0.474				
Downward mobility		0.526				