Which parameters determine farm development in Germany?

Norbert Röder , Stefan Kilian

TUM (Technische Universität München)
Environmental Economics and Agricultural Policy Group
Alte Akademie 14, D-85350 Freising, Germany

Contact: Roeder@wzw.tum.de

European Association of
Agricultural Economists

Università
degli Studi della Tuscia

Istituto nazionale
di Economia Agraria

Rete di informazione
della Commissione Europea

Abstract

In 2005, Germany implemented the Single Payment Scheme which lead to the conversion of direct payments into tradable, production decoupled, single farm payments. The transition from coupled to decoupled support instruments may impact the rate of structural change. The rate of structural change may accelerate since farms with a high share of income derived from CAP payments will abandon farming and lease their land. However, there are also good reasons why the rate of structural change might decrease especially if farmers do not behave as profit maximizers. In Germany agricultural land use is very heterogeneous with respect to management orientation and productivity even at local level. Most of the concerns related to structural change and development of land use intensity, e.g. abandonment of high nature value farmland, are only relevant in a very specific local context. Therefore, it is necessary to establish indicators for farm development on adisaggregated level.

The objective of this paper is twofold. First, we derive criteria and threshold values to classify regions according to their respective natural, socio economic conditions and land use. Second, we evaluate the stability of the link between a set of explanatory variables and the rate of structural change at different spatial scales. Our results indicate that only for a few variables a generally valid link between them and the rate of structural change can be established. For the majority of the explanatory variables, their respective impact on structural change depends heavily on the regional context.

**Key Words:** Structural change, Data mining, Fischler Reform

JEL Code: Q16, Q15, R14

Introduction

In recent years many studies have highlighted the problem of the abandonment of agricultural activities in marginal areas across Europe (for an overview see CABALLERO et al., 2007). This retreat of agriculture is not unproblematic both from the point of view of nature conservancy and of rural development. In Europe, areas of high nature value are concentrated in marginal areas and often associated with low input forms of agriculture (BROUWER et al., 1997), particularly low input grasslands (BIGNAL & MCCRACKEN, 1996). The latter are an important habitat for many endangered species. Empirical studies on rural development from the US suggest that the closure of farms in rural areas does not only affect the agricultural sector itself, but is positively correlated with the rate of emigration of the non-farming population (TWEETEN, 1984).

The marginality of a given region is influenced by four different domains. Environmental, demographic and economic factors, as well as the inhabitants perception of their state determine the relative marginality of a given region (BERTAGLIA et al., 2007; Fig. 1). The environmental factors influence the competitiveness and productivity of economic sectors strongly linked to the environment (e.g. agriculture, forestry, some forms of tourism). Demographic factors can be viewed as a long term indicator of economic prosperity, whereas population density is a good indicator of how rural a region is and the availability of public services.

Fig. 1 - Concept of relative marginality
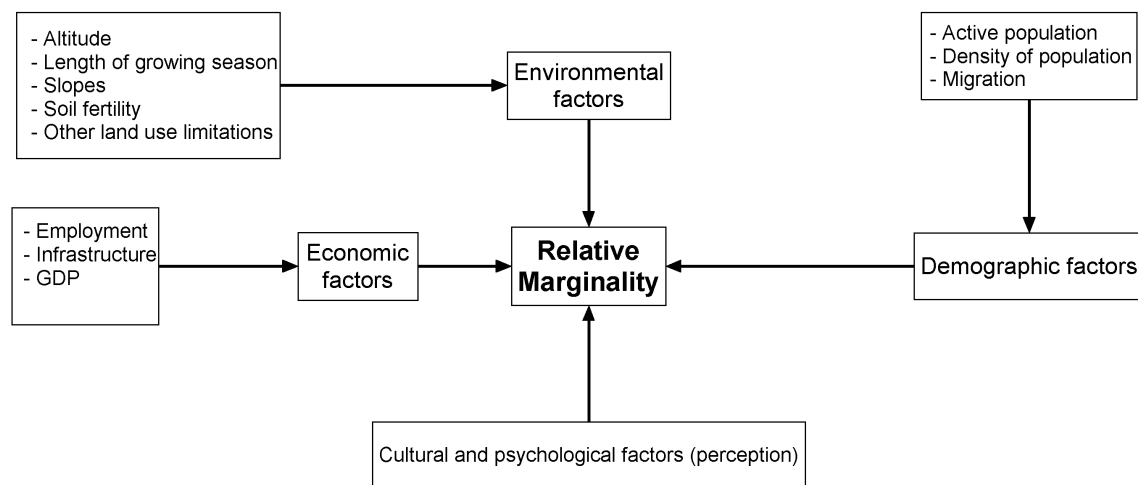
Source: *BERTAGLIA et al. (2007.)*

Two different approaches for the empirical analyses of structural change are commonly used. Either data are analysed on an aggregate level, preferably on a county level, or time series data for individual farms. The first approach has the advantages of longer time series and wider geographic coverage, whereas individual data allow the coverage of farm specific differences (e.g. age, years in business, education, farm history). Unfortunately,

individual longitudinal data are only available in a few countries and are not necessarily sampled in statistically representative fashion (e.g. if based on Farm Accountancy Data Network (FADN) data).

Relevant data in the context of agriculture generally show a high degree of spatial autocorrelation, correlation and at least partial dependencies. This can lead to global regression models, i.e. models incorporating all data in a sample, which are inappropriately specified. One cause of this misspecification are effects which express themselves only if certain thresholds limits for one or more variables are exceeded. This problem is widely recognized in landscape ecology, geography and agronomy and generally handled by a priori stratification of the data, and the calculation of one individual model per stratum. In agricultural economics, a different approach is frequently used. Here, the impact of different strata (regions) is depicted by the inclusion of dummy variables into the global models (e.g. most papers in the following literature review). However, WEISS (2006) shows in a seminal paper for Austria that the direction and the magnitude of impact, a certain explanatory variable has on the structural change, are not only dependent on the value of the variable itself, but frequently depend on the level of other explanatory variables.

A common approach to create more or less homogenous subgroups is clustering. The basic concept behind clustering approaches is the following (cf. WITTEN & FRANK, 2005). Entities belonging to different clusters have barely any features and combination of features in common and therefore they likely belong to different populations. In this case a single global model might yield flawed results and an analysis of clusters one after the other might give largely different insights. However, the application of clustering techniques to stratify agricultural regions for the analysis of agri-economic questions is in most cases linked to rural development and marginal areas (e.g. BERTAGLIA et al., 2007; DAX & HOVORKA, 2005, PFLIMLIN et al., 2005; HELLER, 1997; WÜRFL et al., 1984).

This paper analyses the rate of structural change in Germany from 1999 to 2007. For this cross section analysis is use data aggregated at the municipality level. We estimate a regression in two different settings. In the first one, we estimate one model incorporating all German municipalities. In the second setting the municipalities are partitioned into homogenous groups. For each group a separate model is estimated.

The paper is structured as follows. The following section gives a review of recent literature relating to structural change. Next, we describe the material and applied methods for data manipulation, cluster analysis and regression analysis followed by the presentation of the empirical results. The discussion of the results and the methodological issues finish the paper.

Structural change in agriculture

In recent years many studies were conducted which determine the drivers of structural change in various parts of the developed world. The following section summarizes the key findings of various studies. For a more detailed overview see e. g. MANN (2003) or GLAUBEN et al. (2006).

Many studies across Europe and North America confirm that the rate of structural change, respectively the likelihood of farm exit, declines as farms get larger (e. g. GLAUBEN et al., 2006; HOOPE & KORB, 2006; JUVANCIC; 2006; WEISS, 2006; PIETOLA et al., 2003; HOFER, 2002; BAUR, 1999; KIMHI & BOLLMAN; 1999; WEISS, 1999). Good proxies for the assessment of farm exit rates are the farmer's age and the recent development (e. g. amount of recently rented land, recent investments) of the farm. The likelihood of farm exit is positively related to the farmer's age, as farms frequently close when the owner retires (e. g. GLAUBEN et al., 2006; HOOPE & KORB, 2006; HOFER, 2002; BAUR, 1999; WEISS, 1999). Regarding the history of the farm, farms which were successful in the past have lower farm exit rates. In Central Europe, farms mainly grow by renting additional land, therefore the negative correlation between farm exit rates and the ratio of rented land is comprehensible (GLAUBEN et al., 2006; BAUR, 1999; HOFER, 2002; WEISS, 2006).

Several studies indicate a stabilizing effect of direct payments (GLAUBEN et al., 2006; HOOPE & KORB, 2006, WEISS, 2006; HOFER, 2002; BARKLEY, 1990). Nevertheless in some studies the effect is fairly small (GLAUBEN et al., 2006; HOOPE & KORB, 2006, BARKLEY, 1990).

Comparing different farm types and sectors of agricultural production, HOOPE & KORB (2006) derive lower exit rates for beef farmers than for those involved in cash cropping or hog fattening. WEISS (2006) states that farms specialized in permanent cultures, hogs or poultry fattening or mixed forage cropping are more likely to give up farming than other types. Furthermore, data indicate a negative correlation between stocking density and exit rates. GLAUBEN et al (2006) report a negative correlation for permanent cultures and for the relative ratio of farms keeping cattle for Western Germany.

Other reported influential factors have a less clear connection to structural change. Studies from Austria (WEISS; 1997, 1999), Switzerland (BAUR, 1999; HOFER, 2002) and the US (ROE, 1985) report a positive correlation regarding the connection between off-farm employment and structural change. However, regarding this relation the results of GOETZ & DEBERTIN (2001) and HOOPE & KORB (2006) for the US are ambivalent and several studies indicate even a negative correlation for parts of Canada (KIMHI & BOLLMAN, 1999), Israel (KIMHI, 2000), Western Germany (GLAUBEN et al., 2006) and Slovenia (JUVANCIC, 2006).

For indicators describing the marginality of certain areas in demographic terms, the picture is generally ambivalent. While JUVANCIC (2006), HOOPE & KORB (2006) and GOETZ & DEBERTIN (2001) report a positive correlation of the exit rates and the population density, the results of GLAUBEN et al. (2006) indicate a negative one. For the US, HOOPE & KORB

(2006) and GOETZ & DEBERTIN (2001) report a negative correlation of the exit rates and the distance of the next metropolitan area, while in Austria outside less favoured areas the distance to larger cities is negatively correlated to the exit rates (WEISS, 2006). In contrast, WEISS (2006), BAUR (1999) and JUVANCIC (2006) report lower exit rates in more marginal areas. In HOFER (2002), exit rates and distance are positively correlated. Regarding the connection between regional unemployment rates and farm exit rates, JUVANCIC (2006) and GLAUBEN et al. (2006) report a positive correlation.

Materials

The analysis is conducted on the level of German municipalities. In the federal state of Lower Saxony we used the data of the "Samtgemeinden" and in Rhineland Palantine of the "Verbandsgemeinden" which are comparable to municipalities. Excluding municipalities without any farms and taking into account municipal reform in the investigated time frame, this results in a sample size of 9270 municipalities. For 21 variables identified in previous studies (see e.g. section 2, and BERTAGLIA et al., 2007) as being relevant for determining structural change and defining marginal areas, we could obtain reliable data on the local level (Tab. 1). Inspired by the work of BERTAGLIA et al. (2007) we grouped the variables which depict three different domains. These domains are the site conditions, the agricultural land use and the general socio-economic environment.

Seven variables depict the site conditions. With increasing *altitude,* the conditions for agricultural production become increasingly adverse as vegetation period gets shorter and the precipitation increases. With increasing *relief* (steeper slopes), the conditions for agriculture become problematic due to increases in labour demands and erosion risk.

Generally speaking, the natural conditions in Germany are favourable for the cultivation of cereals. Specific climatic and edaphic conditions are often the reason other cultures may reach above average shares. We use four variables to differentiate these conditions. Permanent cultures like wine and fruit trees are in Germany concentrated in areas with above-average temperatures (*PermCult_UAA*). Root crops, like potatoes and sugar beet, as well vegetable cultivation and horticulture are linked to light, deep and fertile soils (*Root_UAA*). While high shares of the first two variables indicate more favourable condition than on average the remaining two are linked to relatively unfavourable conditions. Permanent grassland (*GrassUAA*) can mainly be found in areas where at least one of the following conditions is met: high summer precipitation, short vegetation period, high risk of late or early frosts or a high groundwater table. Remnants of the potential natural vegetation like forests, moor- and heathland only cover significant shares of the land where the climatic and edaphic conditions for agriculture are unfavourable. Only in these areas do these marginal land uses become economically superior (*MarginalLand*). The last indicator of this domain differentiates the natural potential of the site for the nutrition of ruminants. Grassland is either a marginal form of land use, just before abandonment, or is highly competitive in case of high

summer rain fall and long vegetation periods. While the first case is associated with low stocking densities, the densities in the latter case are exceptionally high (*RCLU_MFA*).

The agricultural production is covered by 12 variables. Ten of them are used in the cluster analysis and 9 in the linear regression models. The farm size is depicted in monetary terms and area by three variables (*GM_farm*, *GM_farm^2*, *UAA_farm*). The amount of 1st pillar payment is depicted by the average value of a single farm payment (*SFP*) in the municipality. Average gross margin per ha is an indicator for the value added per ha of agricultural land (*GM_UAA*).

The composition of a municipality's livestock production is depicted by five variables. *RCLU_LU* and *MFA_UAA* indicate the relative importance of ruminant based systems compared to pig and poultry, and cash cropping, respectively. *LU_UAA* depicts the overall importance of livestock production in a given region.

The ratio of dairy cattle and fattening bulls to the population of ruminants (*DC_RCLU*) represents the high input ruminant production systems while the low input systems are represented by the respective ratio of suckler cows and sheep (*LRCLU_RCLU*). The share of intensive cash crops (maize, wheat, rape seed, etc.) (*Int_CG_UAA*) reflects arable cropping. The intensity of cropping is widely covered in the first domain (site conditions) and is partly covered by *GM_UAA*.

The last domain reflects the general economic conditions on the local level. The gradient between rural and urban areas is covered by the population density (*Pop_dens*). We use the driving time by car to the next larger city (Oberzentrum) as an indicator for the remoteness of a given municipality (*Dis_city*) *UAA_change* reflects mainly the conversion of agricultural land to housing and construction and can be viewed as proxy for the urban pressure on agricultural land. In Germany the differences in the rate of population changes are strongly linked to the flux of population in and out of a community (*Pop_Change*). These levels themselves are strongly related to the general economic prosperity of a given region.

Tab. 1: Variables used in the cluster and regression models

	Variable	Description	Unit	Source	lus-ter	Regres-sion
	Altitude	Avg. altitude of a municipality	m . a.s.l.	USGTOP O30		
	Relief	s.d. of a municipality's altitude	m . a.s.l.			
Land[c]	*Marginal*	Share of total land covered by forests, moor- and heathland in 2000	%	DeStatis (2006)		
A[c]	*Grass_UA*	Share of UAA covered by permanent grassland in 1999	%			
A[a]	*Root_UA*	Share of UAA covered by root crops, horticulture, ... in 1999	%	DeStatis (2008)		
_UAA[a,c]	*PermCult*	Share of UAA covered by permanent cultures (wine, fruits) in 1999	%			**I**
FA[c]	*RCLU_M*	Ruminant LU per ha of main forage area in 1999	L U / ha			
	SFP[c]	Single farm payments per ha in 2005	€ / ha	ZID		
[c]	*GM_UAA*	SGM per ha (based on 1999 land use and average German SGM for the period 2000-2008)	€ / ha	DeStatis (2008); KTBL (2008)		
,c	*GM_farm[b]*	SGM per farm	€ / ha			
c	*GM_farm²*	Squared GM_farm	€² / ha²			
m[b,c]	*UAA_far*	Average farm size in 1999	h a			
A[c]	*MFA_UA*	Share of UAA used as main forage area in 1999	%			
AA[c]	*Int_CG_U*	Share of UAA covered by high input cash crops (wheat, rape seed, maize, root crops, horticulture, tobacco,...) in 1999	%	DeStatis (2008)		
	LU_UAA[c]	LU per ha of UAA in 1999	L U / ha			**I**
U[c]	*RCLU_L*	Ratio of ruminant LU to total LU in 1999	%			
CLU[c]	*LRCLU_R*	Ratio of suckler cow and sheep LU to ruminant LU in 1999	%			
U[c]	*DC_RCL*	Ratio of dairy cow LU to ruminant LU in 1999	%			
	Pop_Cha	Change in the population	%	DeStatis		**I**

nge[c]	between 1996 and 2004		(2006)			
UAA_Cha nge[c]	Change in agricultural land between 1996 and 2004	%	DeStatis (2006)			
Dis_City	Driving distance to the next larger city (Oberzentrum)	m inutes	BBR (2008)			
Pop_Dens d	Population density in 2000	I nh. / km²	DeStatis (2006)			
Stru_Cha nge[c]	Rate of structural change between 1999 and 2007	%	DeStatis (2008)			**D**

Source: Own presentation based on various data sources
a: data are log-transformed (ln(%+0.01)
b: data are log transformed (ln()), transformation of GM_farm and GM_farm² only for the cluster analysis
c: stabilized by a moving window approach
d: data are root transformed
LU: livestock units; UAA: utilized agricultural area; SGM: Standard Gross margin; Inh.: Inhabitants
S: Site; A: Agriculture; G: General
I: Independent variable; D: dependent variable

The average annual rate of structural change with respect to the number of agricultural holdings between 1999 and 2007 is the derived variable for the regression analysis (1).

$$(1)\ c_m := \frac{x_{m,2007} - x_{m,1999}}{8 * x_{m,1999}}$$

c_m: annual rate of structural change

x_m: number of farms in municipality m

m: municipality

Methods

Cluster analysis

We conduct a cluster analysis to define regions with homogenous conditions for agricultural production. Since the municipalities within a given region should be more or less homogenous with respect to all the considered variables, we opt for KMeans, which is a simple partitioning approach (cf. WITTEN & FRANK, 2005). All variables are z-transformed to compensate for the different respective scales and variances. For each of the three data domains we separately conducted a principal component analyses (PCA). With the help of the PCA the joint impact of correlating variables within each domain on the outcome of the cluster analysis is reduced. To account for the different number of variables per domain we standardize the cumulative variance of each of the three PCA to one unit. We use all the axis of the three PCAs as input for the cluster analysis. We determine the number of clusters based on range of criteria in particular r², Akaike and Bayes information criterion (AIC, BIC), Davis-Bouldin's Cluster Validity Indexes (VI43, VI32 (nomenclature according to BASHIRAHAMA (2006))). Finally, we validated the derived clusters with the help of external experts.

Linear Regression

We calculate five different OLS regression models in order to analyze the impact of the explanatory variables on structural change (Tab. 2). In three of these models (A-C) the coefficients are derived for the entire data set (global models) while in the remaining two (D, E), for each cluster an individual OLS is calculated (nested models). In model A the OLS calculation is performed without using regional dummies while model B takes into account only the impact of the regional dummies. Model C combines the previous two models. In model D for each cluster an independent OLS regression is calculated. Model E is based on this model. However, in each sub-model all variables are successively eliminated whose p-value exceeds 0.05.

Stabilization of the data

Between 1999 and 2007 the number of farms in Germany dropped by nearly 97 500 or on average 2.6 % per year (DeStatis, 2008). Taking into account the roughly 10 000 municipalities, this means that in an average municipality about 10 farms closed down in the observed period. However, in 1999, in half of the German municipalities less than 25 farms existed. As a consequence, the closure of a single farm has a large impact on the observed rate of structural change on the municipality level. We apply an approach based on the "moving window" technique in order to stabilize the variables related to the agricultural sector on the municipality level. Actually, the value of each variable assigned to a municipality is a weighted sum of the respective value originally observed this municipality and the values observed in its neighbours (2). If the variable is a relative value the value is derived after the nominator and the denominator are separately derived using formula (2):

$$(2)\ i_{new,m} := 0.5 * i_{old,m} + \sum_n 0.5 / l_n * i_{old,n}$$

i: variable

$i_{new,m}$: derived value of variable i in municipality m

$i_{org,m}$: originally observed value of variable i in municipality m

n: municipality neighbouring m; n neighbours m if n and m share a common border; $n \neq m$

l: number of neighbours of a municipality

For the analysis we take further steps in order to stabilize the data. We transformed some variables having an extremely skewed distribution (*Pop_Dens*, *UAA_farm*, *GM_farm*, *Root_UAA*, *PermCult_UAA*) to spread their data well over their respective range. Furthermore, all extreme values of all variables are truncated to the 0.5% or 99.5% quantils of the respective variable. This step is extremely important for the clustering step, since we use a simple portioning algorithm. If the data would not be well spread over the range of the respective variable, the variable would barely affect the result of the clustering.

Evaluation of the results

We use two indicators analogous to the analysis of association rules to determine the general validity and applicability of the statistically determined correlation (Witten & Frank, 2005). These are the support and the confidence. The support measures the relative number of sub-models in which a certain variable has a significant impact on structural change (3):

$$(3)\ S_i := \frac{Q_i}{T}$$

S_i: Support for variable i

Q_i: number of models in which the variable i is significant at the 5% level

T: number of models

The confidence analyzes how stable the result is for the various sub-models if and only if the respective variable has a significant impact. For the analysis we concentrate on the predominant sign of the respective correlations (4):

$$(4)\ C_i := \frac{|P_i - N_i|}{Q_i}$$

C_i: Confidence of the sign for variable i

P_i: number of models in which the impact is significantly positive at the 5% level

N_i: number of models in which the impact is significantly negative at the 5% level

We calculate each variable's impact to assess the relevance of a given explanatory variable for structural change in the different models (5). This is necessary since the mean and the variance of a given variable differs between each of the sub-models. As a result, the coefficient themselves are barely comparable. Since we are only concerned with the magnitude of the observed impact and not its sign the absolute value is used for calculation.

$$(5)\ I_{t,i} := \left| \sigma_{t,i} * \alpha_{t,i} \right|$$

$\sigma_{t,i}$: s. d. of the variable i in model t

$\alpha_{t,i}$: coefficient of variable i in model t

The average impact of a variable is its weighted impact in the different sub-models (6). For this analysis all coefficients in all sub-models irrespective of their respective level of confidence are considered.

$$(6)\ \bar{I}_i := \frac{\sum I_{t,i} * a_t}{A}$$

$\bar{I}_i$: avg. impact of variable i

a_t: number of observations in model t

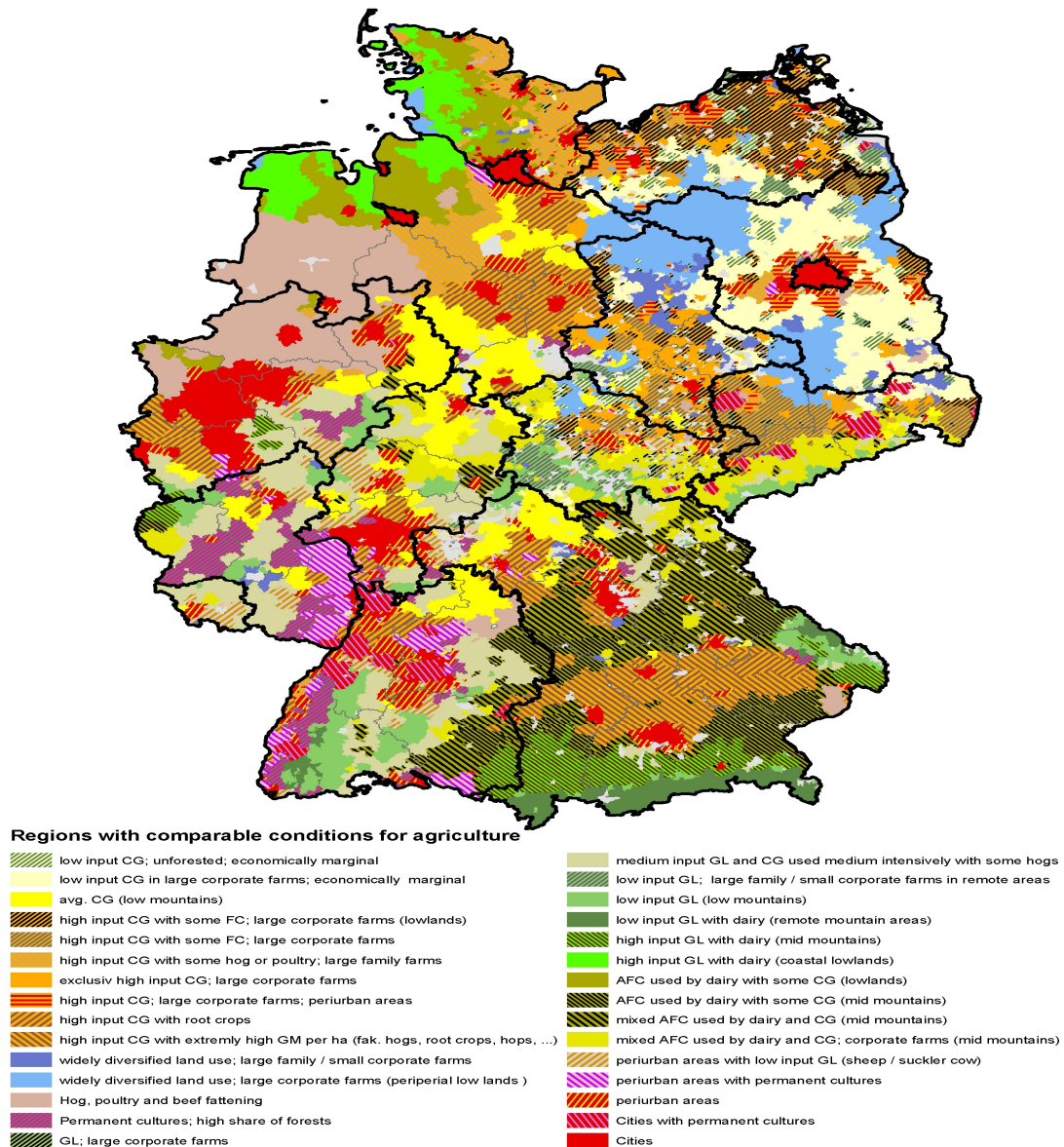A: total number of observations; $A := \sum_t a_t$

We use the following software for the analysis: ESRI ArcMap 9.2 for processing of the geographic data; RapidMiner 4.2 for PCA and cluster analysis; SAS 9.1 for regression analysis; MS Access 2002 for storing and manipulating the data.

Empirical results

In the next section we present the result of the cluster analysis, while the remainder of the paper is devoted to the result of the regression analysis.

Cluster analysis

Fig. 2 - Clusters of homogenous conditions for agricultural production

Regions with comparable conditions for agriculture

- low input CG; unforested; economically marginal
- low input CG in large corporate farms; economically marginal
- avg. CG (low mountains)
- high input CG with some FC; large corporate farms (lowlands)
- high input CG with some FC; large corporate farms
- high input CG with some hog or poultry; large family farms
- exclusiv high input CG; large corporate farms
- high input CG; large corporate farms; periurban areas
- high input CG with root crops
- high input CG with extremly high GM per ha (fak. hogs, root crops, hops, ...)
- widely diversified land use; large family / small corporate farms
- widely diversified land use; large corporate farms (periperial low lands)
- Hog, poultry and beef fattening
- Permanent cultures; high share of forests
- GL; large corporate farms

- medium input GL and CG used medium intensively with some hogs
- low input GL; large family / small corporate farms in remote areas
- low input GL (low mountains)
- low input GL with dairy (remote mountain areas)
- high input GL with dairy (mid mountains)
- high input GL with dairy (coastal lowlands)
- AFC used by dairy with some CG (lowlands)
- AFC used by dairy with some CG (mid mountains)
- mixed AFC used by dairy and CG (mid mountains)
- mixed AFC used by dairy and CG; corporate farms (mid mountains)
- periurban areas with low input GL (sheep / suckler cow)
- periurban areas with permanent cultures
- periurban areas
- Cities with permanent cultures
- Cities

Source: Own presentation CG: cash crops; FC: forage cropping; GM: Gross margin; GL: grass land; AFC: arable forage cropping.

The evaluation of the output of the Kmeans algorithm by the criteria listed in section 4,1 indicated a statistically optimal number of clusters somewhere in the magnitude of 200 and 300. Until this number the information criteria (r^2, AIC, BIC) improve continuously with the number of clusters, while the indicators evaluating the cluster quality (DB- indices) fluctuate within a very narrow range but show an optimum above 200 clusters. Since 200 or more clusters can not be reasonably interpreted, we determine the number of clusters used in the follow up analysis differently. We investigated the cluster models with roughly 30 clusters more in detail since above roughly 30 to 35 clusters the additional information gain per cluster drops sharply. However, it is still larger than expected by chance. Finally, we select a model with 30 clusters based on the overall interpretability of the cluster. In this model the clusters explain 57% of the total variance in the untransformed raw data.

The clusters clearly differentiate the agricultural land use in Germany (Fig. 2). The differentiation into urban, periurban, rural and peripheral areas is clearly visible. Furthermore, the grassland dominated areas along the North Sea and in the mountain areas of central and southern Germany are apparent. In addition the difference in farm size between Eastern and Western Germany is clearly indicated. Also more localized patterns such as the cultivation of fruit trees and wine in the "Alte Land" near Hamburg, at Lake Constance and at the tributaries of the Rhine are depicted.

5.2. Regression models

The different models explain between 27% and 51% of the observed structural change (Tab. 2). The model with regional dummies performs worst (B) while model D performs best. Looking at the BIC the models A, C and E perform equally well while model D is over specified.

Tab. 2: Summary of the results of the regression models

Model	Type	Variables (excl. intercepts)	thereof p < 0.05	R²	Adj. R²	AIC	BIC
A (variables only)	global	18	14	38.2%	38.1%	-5.36	-5.38
B (regions only)	global	30	26	27.1%	26.8%	-5.21	-5.18
C (variables and regional dummies)	global	46	29	40.3%	40.0%	-5.40	-5.37
D (one model per region; all variables)	nested	540	219	51.3%	48.2%	-5.50	-5.08
E (one model per region; variables whose p > 0.05 eliminated by backward elimination)	nested	219	219	49.4%	48.2%	-5.53	-5.36

Source: Own calculation
AIC: Akaike Information Criterion
BIC: Bayes Information Criterion (=Schwarz Information Criterion).

Errore. L'origine riferimento non è stata trovata. depicts the level of support and confidence for the different variables in the reduced nested model (E). The level of support reaches up to 90% for the farm_size (*GM_farm*), i.e. the variable has in 27 of 30 regions a significant impact on structural change. The farm size (*GM_farm*) is always negatively correlated with structural change – if it has a significant influence at all. Therefore it has a confidence of 100%. On the other hand, the impact of the share of grassland (*Grass_UAA*) is highly ambivalent. In 6 regions the correlation to structural change is positive while it is negative in four. This results in a low confidence. Stocking density, the share of marginal land on the total area and the share of ruminants on the total farm animal stock are in many regions negatively correlated with the rate of structural change. The impact of farm size is negatively related to structural change, however, its influence declines as farms gets larger on average. The gross margin per ha (*GM_UAA*) has in many regions a significant influence on structural change. If the influence is significant it is always positively correlated with the dependent variable. Only in a restricted number of regions are *Relief*, *Pop_change*, *SFP*, *PermCult_UAA* and *Pop_dens* of some importance for structural change. While the first three are generally positively correlated with the rate of structural change, the correlation for the remaining two is in most cases negative.

The impact of *Dis_city*, *Root_UAA*, *Grass_UAA*, *Altitude* and *UAA_change* is ambivalent and depends on the given region.

Fig. 3: Support and confidence of the independent variables in the nested model (E)
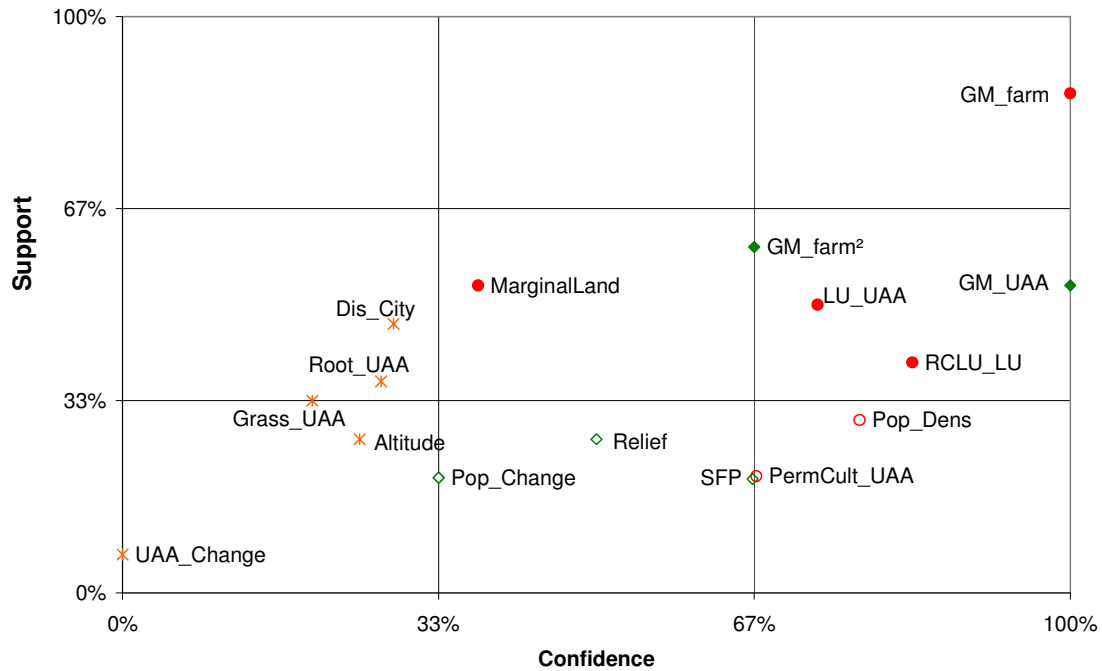
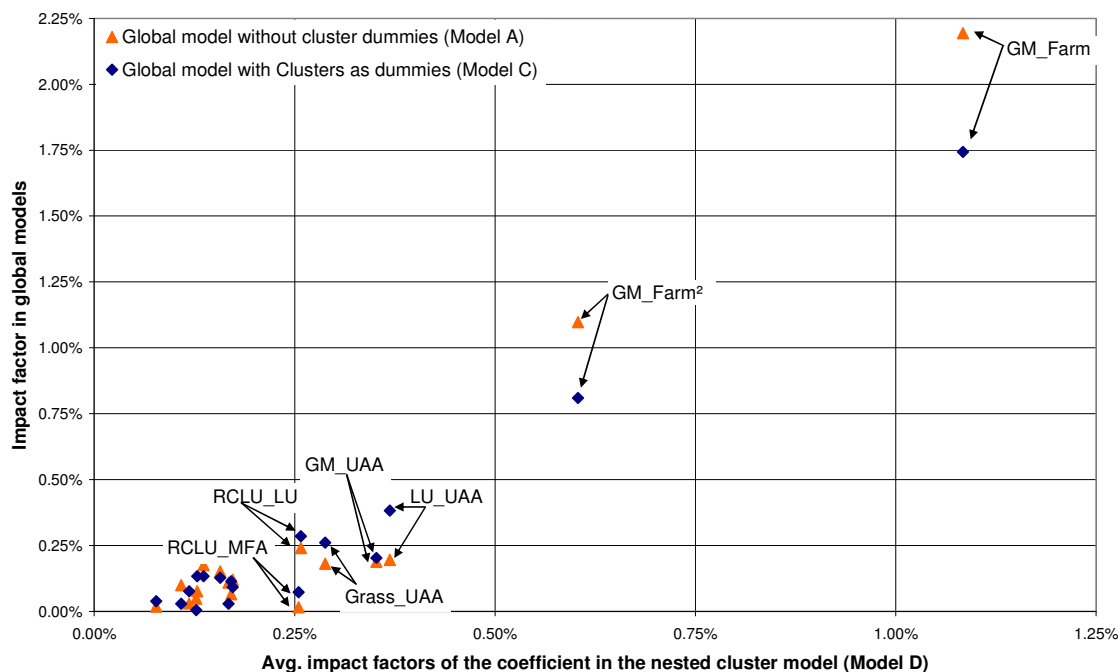*Source: Own presentation based on the analysis of the regression models*
Cubicles: positive correlation between the variable and the rate of structural change predominant
Circles: negative correlation between the variable and the rate of structural change predominant
Filled Cubicles / filled circles: variables having a support exceeding $^1/_3$.

Fig. 4 compares the impact of the different variables in the global and nested models. Farm size (*GM_farm*, *GM_farm²*) is the most important factor in all models. The relationship between farm size and structural change shows that the impact declines and the relationship gets flatter the more regionalized the models become (A→C→D). In contrast the impact gets larger for a group of variables related to livestock husbandry (*RCLU_MFA*, *Grass_UAA*, *LU_UAA*) and the intensity of land use (*GM_UAA*) as the models become more regionally differentiated. *RCLU_LU* is the only parameter that is of overall relevance and whose is impact is fairly independent of the model specification. The remaining variables have in common that their respective impact is fairly low irrespectively of the chosen model specification

Fig. 4: **Comparison of the magnitude of the impact of the independent variables in the nested model (D) and the global models (A, C)**

Source: Own presentation based on the analysis of the regression models.

Summary and conclusions

If one looks more closely at the regions in which a designated variable becomes significant, a fairly consistent picture emerges for many of the variables. Farm size, in its linear transformation, is only indifferent in areas dominated by large corporate farms. Here, structural change can be better explained by a negative correlation to the squared farm size. In all other regions, the rate of structural change is negatively correlated to the linearly transformed farm size. However, its impact declines as farms get larger. The altitude is only important in the transition from the lowlands to the lower mid mountain areas and in the higher mid mountain and mountain areas. An increase in the relief energy leads only to a higher rate of structural change if the terrain is more or less flat; while it has no effect in areas with a more undulated terrain. In most cases when the share of grassland is important for structural change the agricultural land is mainly devoted to crops or permanent cultures. Only in areas with a high share of permanent cultures the rate of structural change declines as their share increases. At low shares an increasing share of root crops is associated with a declining rate of structural change. However, at high densities the correlation reverses. Structural change is only positively correlated with the gross margins per ha in regions where the gross margin is above the German average. The rate of structural change is the highest in regions with an intermediate stocking density, while it declines as stocking densities gets higher or lower. The ratio of ruminants to the total animal stock only has an impact on structural change at intermediate levels.

While in peripheral areas the rate of structural change increases the more remote the area gets, the rate of structural change declines as one move from urban areas to rural areas in their vicinity. The population density has a negative impact on the rate of structural change only in more or less urban areas.

In general, the sub-models have a higher predictive power in West Germany than in East Germany and in regions with a higher productivity of the agricultural land. The correlation of variables related to the productivity per farm or unit of area (*GM_farm*, *GM_farm², GM_UAA, MarginalLand)* or to the intensity of type of livestock husbandry (*LU_UAA, RCLU_LU*) show a high confidence and support. The other variables have either only in a few regions a relevance effect or the sign of the effect depends on the regional characteristics. Therefore, one can not transfer results from one region to another one to one.

In order to improve the confidence in the obtained results, different cluster algorithms (e.g. expectation maximization, density based) should be applied to the data set. Furthermore, the impact of different forms of data manipulation (aggregation, transformation and truncation) on the clusters should be tested. Furthermore, the introduction of variables depicting the distribution of a given variable on a local scale might yield additional insights (e.g. based on the coefficient of variation, the Gini coefficient or the Shannon-Weaver Index).

In Germany, grassland based ruminant production systems are fairly important in rural and marginal areas. In the areas dominated by these types of farming the rate of structural change was in the recent years on average lower than in regions dominated by more intensive production systems or cropland. Therefore, one may conclude that, currently the risk of an abandonment of agriculture in these areas and consequently putting at stake the natural values (depending on low input types of agriculture) is comparatively low.

References

Barkley, A. P. (1990): The Determinants of the Migration of Labour out of Agriculture in the United States, 1940 - 85. American Journal of Agricultural Economics 72: 567-573.

Baur, P. T (1999): Agrarstrukturwandel in der Schweiz. Dissertation ETH. Nr. 13240. Zürich.

BBR (Federal Office for buildung and regional planning) (2008): Erreichbarkeitsindikatoren. personal communication.

Bertaglia, M., Joost, S., Roosen, J. & Econogene Consortium (2007): Identifying European marginal areas in the context of local sheep and goat breeds conservation: A geographic information system approach. Agricultural Systems 94: 657-670.

Bignal, E. M. & McCracken, D. I. (1996): Low-intensity farming systems in the conservation of the countryside. Journal of Applied Ecology 33: 413-424.

Brouwer, F., Baldock, D., Godeschalk, F. & Beaufoy, G. (1997): Marginalisation of agricultural land in Europe. In "Livestock in European Rural Development" (L. Laker & J. Milne, eds.): 25-32. Macaulay Land Use Research Institute. Aberdeen, Scotland.

Caballero, R., Riseth, J. A., Labba, N., Tyran, E., Musial, W., Molik, E., Boltshauser, A., Hofstetter, P., Gueydon, A., Röder, N., Hoffmann, H., Moreira, M., Seita Coelho, I., Brito, O., & Gil, A. (2007): Comparative Typology in six European Low-intensity Systems of Grassland management. Advances in Agronomy 96: 351-420.

Dax, T. & Hovorka, G. (2004): Berggebiete in Europa - Ergebnisse des internationalen Forschungsprojektes zur Abgrenzung. Situation und Politikanalyse. Bundesanstalt f. Bergbauernfragen. Facts & Features 32, p. 38. Wien.

DeStatis (Federal Statistical Office) (2006): Statistik lokal (cadastre & population census). (CD). Wiesbaden.

DeStatis (Federal Statistical Office) (2008): agricultural structural surveys. scientific access. various years.

Goetz, S. J. & Debertin, D. L. (2001): Why Farmers Quit: A County-Level Analysis. American Journal of Agricultural Economics 83: 1010-1023.

Heller, A. (1997): Neue Typen der Agrarstruktur Österreichs. PhD- Thesis. University of Innsbruck. p. 189. Innsbruck.

Hofer, F. (2002): Strukturwirkungen von Direktzahlungen. Dissertation ETH. Nr. 14464. Zürich.

Juvancic, L. (2006): Determinants of Farm Survival and Growth in Slowenia (1991-2000). Paper presented at the 96th seminar of the EAAE 'Causes and Impacts of Agricultural Structures' from the 10-11.01.2006. p. 17. Tänikon.

Kimhi, A. (2000): Is Part-time Farming Really a Step in the Way Out of Agriculture? American Journal of Agricultural Economics 82: 38-48.

Kimhi, A. & Bollman R. (1999): Family Farm Dynamics in Canada and Israel: The Case of Farm Exits. Journal of Agricultural Economics 21: 69-79.

KTBL (Kuratorium für Tecknik und Bauwesen in der Landwirtschaft) (2008): Database for Standard Gross margin; scientific access, lastly accessed 01/0872008; URL: http://daten.ktbl.de/sdb.

Mann, S. (2003): Theorie und Empirie agrarstrukturellen Wandels? Agrarwirtschaft 52: 140-148.

Momin, B. F. (2006): Clustering and Validation for Very Large Databases (VLDB). International Conference on Information and Automation (ICIA): 258-263.

Pflimlin A., Buczinski B. and Perrot, C. (2005): Proposition de zonage pour préserver la diversité des systèmes d'élevage et des territoires européens. Fourrages 182: 311-330.

Pietola, K., Vare, M. & Lansink, A. O. (2003): Timing and Type of Exit from Farming: Farmers' Early Retirement Programmes in Finland. European Review of Agricultural Economics 30: 99-116.

Roe, B. (1995): A Study of US Farm Exits with Evidence from the Panel Study of Income Dynamics. Paper presented at the 1995 Annual meeting of the American Agricultural Economics Association.

Tweeten, L. (1984): Causes and Consequences of Structural Change in the Farming Industry. Planning Report #207. Washington DC.

Weiss, C. R. (1997): Do They Ever Come Back Again? The Symmetry and Reversibility of Off-Farm Employment. European Review of Agricultural Economics 24: 65-84.

Weiss, C. R. (1999): Zum Ausscheiden landwirtschaftlicher Betriebe: Eine Empirische Analyse, Agrarwirtschaft, 48/5, 202-209.

Weiss, F. (2006): Bestimmungsgründe für die Aufgabe - Weiterführung landwirtschaftlicher Betriebe in Österreich. Univ. für Bodenkultur, Inst. für Nachhaltige Wirtschaftsentwicklung. discussion paper DP-14-2006. p.32. Wien.

Witten, I. H. and Frank E. (2005): Data Mining: Practical Machine Learning Tools and Techniques.

Würfl, P., Dörfler, J. & Rintelen, P. M. (1984): Die Einteilung Bayerns in landwirtschaftliche Standorte, Landwirtschaftliche Erzeugungsgebiete und Agrargebiete. Bayerisches Landwirtschaftliches Jahrbuch 61: 377 – 396. München.

USGSGTOPO30 (U.S. Geological Survey Center for Earth Resources Observation and Science (EROS)) (2006): Digital elevation model, Sioux Falls.