INSTITUTE
OF ECONOMIC STUDIES
Faculty of Social Sciences
Charles University

REGIONAL CONVERGENCE IN THE EUROPEAN UNION: WHAT ARE THE FACTORS OF GROWTH?

Jan Pintera

IES Working Paper 20/2021

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Pintera J. (2021): "Regional Convergence in the European Union: What are the Factors of Growth?" IES Working Papers 20/2021. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

Regional Convergence in the European Union: What are the Factors of Growth?

Jan Pintera

Institute of Economic Studies, Faculty of Social Sciences, Charles University
Prague, Czech Republic, email: jan.pintera@fsv.cuni.cz

June 2021

Abstract:
Despite years of deepening economic integration among the states and regions of the European Union, empirical research remains inconclusive about speed of convergence across regions, if not its existence. This paper provides a new look on convergence in the EU while focusing on development at regional level after the Great Recession. It uses the log t convergence test by Phillips and Sul (2007) to analyze the convergence in level of income among the European regions. Rather than supporting the convergence hypothesis, we identify five convergence clubs in which the regions converge in income growth rates. Investigating further the geographical distribution of the convergence clubs, we confirm high inequality within the member states and find large continuous area of high convergence clubs in the urbanized part of Western Europe. Furthermore, we investigated the determinants of convergence club membership using Logistic Regression. We found a low impact of any of the estimated variables on membership in the highest club but confirmed positive association of membership in the higher clubs with research and patent activities.

JEL: C23, C40, R11
Keywords: Club Convergence, European Regions, log t test, Logistic regression

1 Introduction

The aim of this work is to provide evidence on income convergence across the regions of the European Union after the Great Recession. Convergence in standard of living among the regions of the European Union can promote socio-economic homogeneity and thus pave the way for political consensus. What makes convergence an attractive from the viewpoint of empirical economics, is how complicated relationship can be found during the history of the European integration at the regional level with many instances of persistent (or even rising) differences between certain regions and the speed and character of convergence changing over time (Eckey and Türck, 2007; Zarotiadis and Gkagka, 2013; Iammarino et al., 2019).

Previous empirical research has suggested several distinguishing features of convergence in the EU. It is often pointed out that economic development in large urban areas, and above all capital cities, is found significantly more dynamic than in the rest of the countries, leading to high economic inequality within the EU states (Geppert and Stephan, 2008). This seems to be especially valid for the new member states with capital cities and main urban regions being already among the EU's richest regions by GDP per capita, while many other regions lagging behind and the overall discrepancy between the old and the new member states diminishing rather slowly. The result is that internal inequality in the CEE countries is considered especially high (Geppert and Stephan, 2008; Smetkowski and Wójcik, 2012; Smętkowski, 2013).

This work investigates the dynamics of regional income convergence with the main focus on finding out whether there is a convergence among all the EU countries or there is evidence of convergence clubs. The work uses the Phillips and Sul's test that is relatively widely used to analyze convergence at various levels (Bartkowska and Riedl, 2012; Borsi and Metiu, 2015; Von Lyncker and Thoennessen, 2017) and its design is especially suited for investigation of club convergence that can be expected among EU regions. Furthermore, we employ the logistic regression on the resulting convergence clubs, examining the contribution of wide range of variables, suggested in the economic growth literature, to the current state of convergence in the EU.

Our results imply the club convergence hypothesis, rather than overall convergence among all regions. In contrast to Cutrini (2019), who uses a similar methodology, our results support existence of large continuous area of high convergence regions in the urbanized part of Western Europe (the "Blue

Banana"), rather than an existence of any Central European manufacturing core encompassing regions of both old and new EU member states.

This work will thus contribute to the existing literature in two ways. First, it will examine the economic convergence in the EU in the post-crisis environment and second, it will attempt to draw a connection between various socio-economic characteristics of the EU regions and their relative economic performance.

The paper is structured as follows: section 2 provides an overview of the methods used for convergence analysis. Section 3 describes the Phillips and Sul's test and convergence clubs formation procedure in detail. In Section 4 we discuss the test results and convergence clubs composition. Section 5 analyses the factors determining convergence club membership and finally, section 6 provides conclusions and policy implications.

2 Literature review

2.1 The convergence hypothesis and modelling frameworks for its empirical measurement

Eckey and Türck (2007) bring comprehensive survey of methods used for convergence analysis and their empirical results. Probably the most common form of the convergence test is the β-convergence which departs from the transitional income dynamics equation (3) and in its basic version has the following form:

$$\log y_{it} = a_0 + (1 - a_1) \log y_{it-1} + u_{it} \tag{1}$$

where we assume $0 < a_1 < 1$. In this setting, $a_1 > 0$ implies convergence as the growth rate $\log y_{it} - \log y_{it-1}$ is inversely related to income in period t. The error term u_{it} in the equation captures all sorts of temporary changes in the parameters of the production function.

The regression above is then often augmented by control variables creating the conditional convergence model (Sala-i Martin, 1996). It is worth noting that the β-convergence is designed to test the sign of the convergence parameter β in equation (3) while we assume the parameter's homogeneity across time and cross-sections. This assumption is the main difference between classical β-convergence and the log t test by Phillips and Sul (2007) used in this work.

Other common approach is the so-called σ-convergence that is based on the sample variance of log income per capita which is used in combination with the equation for $\log y_{it}$ above to derive following first-order difference equation: $\sigma_t^2 \cong (1-a_1)^2\sigma_{t-1}^2 + \sigma_u$. This equation shows the relation with β-convergence as σ-convergence convergence is only possible if $0 < a_1 < 1$. However, using the steady state value of $(\sigma^2)^*$ to derive the expression

$$\sigma_t^2 = (\sigma^2)^* + (1-a_1)^2[\sigma_{t-1}^2 - (\sigma^2)^*] \tag{2}$$

and, as noted by Sala-i Martin (1996), β-convergence does not have to imply σ-convergence as it depends on the initial value of σ^2 whether the sample variance σ_t^2 will be increasing or decreasing on its way to the steady state.

The β and σ-convergence are still broadly used frameworks for estimation of convergence, recently, however, many other methods were utilized for finding convergence at both state and regional level. Following part presents a brief description of some of these methods and their results.

2.2 Empirical analyses of convergence among EU regions

Eckey and Türck (2007) conclude in their meta-analysis that most of the empirical convergence literature finds significant, however rather small, convergence rate among European regions in the time span from 1980s to early 2000s. This finding is, however, not universal. Some papers find high rate of convergence, whereas others fail to find convergence at all. This variety of conclusions is caused by differences in methods as well as in number of regions included (Eckey and Türck, 2007).

Some authors also suggest that convergence among the European regions was not stable in the post-war period. Eckey and Türck (2007) note that several empirical works found that convergence rate among the EU regions had diminished during the second half of the 20th century. Other researchers concluded that convergence in Europe appears to be U-shaped (Basile et al., 2001; Geppert and Stephan, 2008). It seems to reach its lowest point around the beginning of the 1980s, accelerating again later on. Geppert and Stephan (2008) also conclude that convergence is happening mainly at the national level whereas regionally we do see strengthening of metropolitan areas. From the newer contributions, Smetkowski and Wójcik (2012) find, using the β-convergence, only a weak tendency for regional convergence in the CEE countries, while in the case of some countries (e.g. Poland) there was absolute divergence even when the capital region was excluded. As far as the σ-convergence is concerned, the results seem as inconclusive as before with conclusions differing based on the period and the sample of regions investigated (Eckey and Türck, 2007).

Another concept of convergence that appears to gain popularity in recent years is club convergence. A standard model for testing this type of convergence is the so-called LISA (Local Indicators of Spatial Association). It is based on finding clusters of regions with value of a variable measuring spatial association between the regions constantly above average for a longer period of time. Statistic defined by Getis and Ord (1992) is then used as this variable most often. One of the distinguishing features of the LISA methodology is that it is aimed at finding clusters of neighbouring regions. In result, this method often leads to large clusters along the south-north (as found for example in Baumont et al. (2003)) or between new and old members states (Eckey and Türck, 2007).

Other approaches used for convergence clubs detection are Kernel density estimation, Markov chains (Eckey and Türck, 2007), Bayesian (Fischer and LeSage, 2015) and clustering methods (Maasoumi and Wang, 2008). Among the methods used for convergence analysis is also the log t test by Phillips and Sul (2007) used in this work. This framework can distinguish between absolute and relative convergence and detect convergence even in situations where traditional convergence tests fail thanks to its treatment of convergence as an asymptotic property (Bartkowska and Riedl, 2012; Borsi and Metiu, 2015). Moreover, as results of the papers cited below suggest, the log t test is able to produce convergence clubs, members of which are not geographical neighbours. This is in contrast with the LISA analysis, described previously, which almost necessarily lead to large continuous blocks of regions. The Phillips and Sul (2007)'s test is relatively widely used at national level, as shown by Borsi and Metiu (2015); Fritsche and Kuzin (2011); Monfort et al. (2013); Apergis et al. (2010) for Europe and by Rodríguez-Benavides et al. (2014) for Latin America. The test is also used at regional level where both Bartkowska and Riedl (2012) and Pinho et al. (2010) work with Western European regions. In the first case between years 1990 and 2005, in the second case between 1980 and 2007, both using NUTS 2 units. Ghosh et al. (2013) employs the same method for states in India. Bartkowska and Riedl (2012) found, using the log t test, 5 separate clubs in Western Europe. Analyzing the spacial distribution of these clubs, they found an agglomeration effect among the Western European regions in form of tendency of the regions with capital cities to belong to higher convergence clubs than their neighbouring regions. Furthermore, there is also a tendency of regions within one country and regions belonging to the same club to cluster together (Bartkowska and Riedl, 2012). The newest contribution for Western Europe is Von Lyncker and Thoennessen (2017), finding 4 convergence Clubs on 1980–2011 regional dataset.

Other empirical works confirm an inclination of large urban areas, and above all capital cities, to grow faster than other regions. Especially often is then mentioned resulting large regional income inequality, particularly strong among the CEE countries (Cuaresma et al. (2014); Smetkowski and Wójcik (2012); Szendi (2013); Chapman et al. (2012); Monastiriotis (2011). Smetkowski and Wójcik (2012) conclude, using the LISA analysis, that in the CEE countries regions with large metropolitan areas and some almost stagnating agricultural areas are forming different convergence clubs, nevertheless there is also relatively high income mobility found in case of the remaining regions. Interestingly, the authors find that the slowly growing regions tend to be often located at the eastern border of the EU or on some geographically disadvantaged locations. Monastiriotis (2011), among others, mentions a strong tendency to growing income inequality in CEEC visible already shortly after the fall of the Iron Curtain. According to Monastiriotis this increase is many times higher than in case of the old EU members. Persistent regional disparities in the CEE countries, even in composite measures of well-being (Human Development Index), are confirmed also by Benedek and Kocziszky (2015). They found no signs of σ-convergence between 1995 and 2000.

2.3 Determinants of regional development

Mora (2008) considers the club convergence as sort of poverty trap, often related to non-convexity of aggregated production function. Specifically, he mentions market characteristics (namely size and structure) as possible factors influencing the club convergence. In line with Galor (1996), he considers several factors that can lead to club convergence based on study of the neoclassical growth model. Especially, differences in initial human capital distribution and in initial income distribution can bring otherwise structurally same economies to different equilibria.

Guastella and Timpano (2016) then contribute to the convergence debate from the viewpoint of New Economic Geography (NEG) and New Growth

Theory (NGT) that are focused on endogenous factors that drive the regional development. This theory, unlike the neoclassical growth model, predicts long-run divergence or conditional convergence in case of NEG and core-periphery divergence in case of NGT. As the drivers of the endogenous development are then considered knowledge spillovers, investment in innovation activities or spatial concentration of economic activities (Guastella and Timpano, 2016).
The club convergence hypothesis is often expressed in form of a regression with multiple control variables such as saving rates, population growth variable, human capital endowment, proxy for R&D (patent application per million inhabitant). Authors also usually include sectoral specialization indexes and other variables (Mora, 2008). Bartkowska and Riedl (2012) consider, among other variables, share of high-tech manufacturing and services in total manufacturing and services and share of services in total GVA, as structural characteristics. Guastella and Timpano (2016) test the above mentioned hypothesis of endogenous growth by using the classical cross-sectional β-convergence framework extended for variables representing the growth factors mentioned above, with Knowledge Intensive Business Services used as a proxy for human capital.

This work follows the above mentioned literature in its choice of variables for logistic regression that represents a key feature in our analysis. The particular choice of explanatory variables is described in more detail below.
One of the novelties of this paper is that we focus on the impact of a specialization in Business services on regional convergence. Business services can be described as activities that enhance efficiency and quality of their customers' production processes. They include among others computer and IT related services (which includes software services), R&D (excluding the university based R&D), engineering and technical consultancy and temporary labour recruitment services and training (Kox and Rubalcaba, 2007; Hertog, 2000). BS without doubts experienced a remarkable growth process in the past decades - they were certainly the fastest growing sectors in terms of employment between 1979 and 2003 (Kox and Rubalcaba, 2007).
Direct contribution of BS to the economic growth is less clear. Even though Desmarchelier et al. (2013) show that KIBS (Knowledge Intensive Business Services, a subset of Business Services, arguably comprising the more innovative parts of BS [1] are a factor of economic growth in a theoretical model

[1] As described by Kox and Rubalcaba (2007), the Business Services are formed of KIBS

framework, BS have also relatively weak productivity performance (Kox and Rubalcaba, 2007). Empirical results on the impact of BS on aggregated productivity and growth, as summarized by Kox and Rubalcaba (2007), show then rather ambiguous picture - there are no robust results confirming positive impact of BS, when taken as a whole. Pylak and Majerek (2014) show, using structural equation modelling (SEM), that KIS (Knowledge-Intensive Services - more general category of service, including knowledge intensive BS) have different impact in highly and less developed regions. Pylak and Majerek (2014) conclude that in case of the less developed regions KIS may have more direct impact on growth due to probable inefficiencies of the overall industry. However, for example Corrocher and Cusmano (2014) find strong relationship between BS and innovation-oriented regions while stating that lack of BS is strongly associated with regions poorly performing in innovation activities. Similar are conclusions of Muller and Zenker (2001).

3 Methodology

3.1 The log t test - modelling framework

According to Phillips and Sul (2007), the main drawback of classical β-convergence is its assumption of homogeneous technology process leading to time and country invariant coefficient β in (3). Phillips and Sul (2007) show that this assumption leads to inconsistency of classical β-convergence stemming from omitted variable bias as well as endogeneity. The log t test by Phillips and Sul (2009) tries to overcome the problems of classical β-convergence and builds a robust and flexible testing procedure capable of capturing multiple types of convergence observed in reality.

The model departs from a variation of the neoclassical growth model where the transitional cross-sectional divergence is possible thanks to the fact that parameters β_{it} and x_{it} are allowed to vary across cross-sections:

$$y_{it} = y_i^* + a_{i0} + (y_{i0} - y_i^*)e^{-\beta_{it}t} + x_{it}t \tag{3}$$

Variables y_{i0} and y_i^* in (3) are initial and steady state levels of log per capita income, x_{it} expresses technology accumulation over time and a_{i0} cap-

and Operational Business Services. The former include all technical and IT related categories. The latter consist of administration, bookkeeping, security services and various types operational services tures initial technology accumulation. β_{it} is a transition parameter and, together with the technology accumulation parameter x_{it}, it is assumed to be homogeneous across countries in the neoclassical theory.

The authors assume that β_{it} is an increasing function of the technological progress parameter x_{it}. This can explain divergence and income traps among countries (or regions); as the technology parameter x_{it} is idiosyncratic, poor countries can differ in their technological progress from the developed ones, which gives rise to differences in speed of convergence given by parameter β_{it}.

If we assume $x_{it}t$ to contain both idiosyncratic and shared element across economies and we express the equation (3) as

$$y_{it} = (\frac{y_i^* + a_{i0} + (y_{i0} - y_i^*)e^{-\beta_{it}t} + x_{it}t}{\mu_t})\mu_t = \delta_{it}\mu_t \tag{4}$$

Here, μ_t represents common growth component that is being shared among economies in question. μ_t is a broadly defined trend that can have both deterministic and stochastic component and can arise, for example, from knowledge and technology sharing among countries. μ_t also determines common growth in the steady state. δ_{it} then captures how much is particular economy close to steady state growth represented by μ_t.

Note that by defining $a_{it} = y_i^* + a_{i0} + (y_{i0} - y_i^*)e^{-\beta_{it}t}$, we can see that as $t \to \infty$ the time dependent element $(y_{i0} - y_i^*)e^{-\beta_{it}t}$ of a_{it} decays to zero and y_{it} in (3) thus in the long-run starts to follow solely the path of technology accumulation $x_{it}t$. We can therefore think of component μ_t as being dependent on common technology development.

Using previous definition of a_{it}, we can rewrite of loadings as $\delta_{it} = \frac{a_{it}+x_{it}t}{\mu_t}$. Moreover, if we represent the common steady state growth element μ_t by either unit root stochastic trend or simple linear deterministic trend $\mu_t = t$, we can see that

$$\delta_{it} = x_{it} + \frac{a_{it}}{t} \tag{5}$$

and thus $\delta_{it} \to x_i$ as t comes to infinity, assuming that x_{it} converges to x_i. As x_{it} determines steady-state behaviour of y_{it}, as noted above, δ_{it} therefore plays a key role of a transition parameter. Parameter δ_{it} is supposed to have following structure: $\delta_{it} = \delta_i + \sigma_{it}\xi_{it}$ where $\sigma_{it} = \frac{\sigma_i}{log(t)t^{\alpha}}$. The parameter α then sets the rate at which $\delta_{it} \to \delta_i$ with $t \to \infty$ and can be interpreted as the speed of convergence. In particular, the convergence of δ_{it} to δ_i is guaranteed for all $\alpha \geq 0$. This inequality therefore becomes subject of the null hypothesis of the test below together with condition of shared value of δ_i across cross-sections:

$$H_0 : \delta_i = \delta \quad \& \quad \alpha \geq 0 \tag{6}$$

Thus, we test overall relative convergence among the cross sections, with the alternative allowing both overall divergence or club convergence:

$$\begin{aligned} H_A : \{\delta_i = \delta \quad \text{for all i with} \quad \alpha < 0\} \quad \text{or} \\ \{\delta_i \neq \delta \quad \text{for some i with} \quad \alpha \geq 0, \quad \text{or} \quad \alpha < 0\} \end{aligned} \tag{7}$$

For testing procedure as well as modelling of the transition parameter δ_{it} is then used following formula:

$$h_{it} = \frac{y_{it}}{N^{-1} \sum\limits_{i=1}^{N} y_{it}} = \frac{\delta_{it}}{N^{-1} \sum\limits_{i=1}^{N} \delta_{it}} \tag{8}$$

This formula traces the trajectory of each cross-section i relative to the club's average and is thus called by Phillips and Sul (2009) the relative transition path. It also reflects any divergence of individual unit i from the common trend μ_t. While individual transition paths h_{it} may be various, including transitional or permanent divergence, the ultimate growth convergence implies $h_{it} \rightarrow 1$ as $h_{it} = h_t$ in case of ultimate convergence to a common trend.

In the convergence test itself, the authors concentrate on cross-sectional convergence of individual δ_{it}. Mainly due to ease of calculation from the data the authors concentrate on h_t, rather than δ_{it} itself, in the test derivation and they compute mean square cross-sectional "transition differential" of h_{it}:

$$H_t = N^{-1} \sum_{i=1}^{N} (h_{it} - 1)^2 \tag{9}$$

As the ultimate growth convergence implies $h_{it} \rightarrow 1$, the value H_t, which can be also interpreted as a quadratic distance of the club from the common limit, has to converge to zero with time going to infinity. If it remains positive, we conclude that convergence did not happen (Phillips and Sul, 2009).

The test itself is then motivated by the problem that it is relatively hard to distinguish whether H_t converges to zero or to a constant. Phillips and Sul (2007) therefore developed a model based on following OLS regression, together with a testing procedure introduced later. They first show that under the model specification shown above the term H_t has following limiting form: $H_t = \frac{A}{log(t)^2 t^{2\alpha}}$ as $t \to \infty$, which then leads to the final formulation of log t regression:

$$\log(\frac{H_1}{H_t}) - 2\log(\log(t)) = a + b\log(t) + u_t \tag{10}$$

The test (called the log t convergence test) is then one sided t-test test of convergence against no or partial convergence. Coefficient b converges in probability to the speed of convergence parameter 2α and the convergence hypothesis is tested by one-sided t-test of inequality $\alpha \geq 0$, using the estimated parameter b with HAC standard errors. In more detail, the t-statistic of the test's parameter converges to either positive infinity in case of $\alpha > 0$ or weakly to a standard normal distribution in case of $\alpha = 0$. Under the alternative the estimate of b converges to zero but the t-statistics diverges to $-\infty$.

The type of convergence can be recognized by the magnitude of the b coefficient, which measures the speed of convergence, namely if $b \geq 2$ (which means $\alpha \geq 1$), we can see level convergence (i.e. we conclude that the regions converge in levels of per capita incomes), whereas if $0 \leq b < 2$, there is a relative convergence, implying convergence only in the income growth rates over time.

3.2 The log t test - formation of convergence clubs

If the hypothesis of overall convergence is rejected we check for convergence in subgroups of the investigated sample. The clustering procedure follows these steps:

- Order the individuals by the amount of last period income (or other variable).
- A core group of k^* highest individuals is chosen by maximizing the log t test's statistic t_k over the various sizes of k^*: $k^* = argmax_k\{t_k\}$ subject to min $\{t_k\} > -1.65$. If $t_k \leq -1.65$ for $k = 2$, the highest individual is dropped and this step repeated starting from the second highest observation.
- One region at a time is added to the core group, formed in the previous step, and the log t test is run again. The respective t-statistics is than

compared to criterion level c^*. In our case we choose $c^* = 0$. If the associated t-statistic is greater than c^*, we add the individual to the club.

- We run the log t test for all the remaining observations, if they fulfill $t_b > -1.65$ we conclude they form a second convergence club. If not, we repeat all the previous steps with the remaining observations, to see whether we can find convergence clubs among these remaining individuals.

$c^* = 0$, in the second step of the procedure plays an important role in final composition of the convergence clubs, with higher values of c^* meaning lower probability of including a wrong region to the club. Phillips and Sul (2009) note that c^* can vary between 0 and -1.65 and $c^* = 0$ is considered as a very conservative choice, which tends to detect larger number of clubs than it should. On the other hand, $c^* = -1.65$ was recommended when we posses relatively large dataset. As our data starts from various reasons only after 2003, we adopt $c^* = 0$ combined with a club merging procedure suggested by Phillips and Sul (2009) and further elaborated by Bartkowska and Riedl (2012).

This procedure contains a step by step merging of several groups together and testing whether the log t test statistics of this merged group is larger than -1.65. If it is larger, we conclude these two clubs form a convergence club together. Concretely, we start by merging first and second club together and proceed by adding following clubs until the null hypothesis of the log t test is rejected. We conclude that all clubs that passed the test form a single convergence club. Subsequently, we continue starting again from the first club for which this merging hypothesis was rejected. We try to merge it with all remaining clubs, using the same procedure. If the null hypothesis is rejected for the first and second club, we leave the first club untouched and start by merging the second with third club and continue in the same fashion as just described.

As often noted (Dall'Erba and Le Gallo (2008), Magrini (2004), Anselin and Rey (1991), Anselin (2001)) spatial autocorrelation is a frequent problem in analysis of regional units, leading to biased t-tests and measures of fit. We indeed confirm existence of autocorrelation among the EU regions in log GPD per capita in our data. This could harm the results of Phillips and Sul's log t test, as the key parameter b is based on an OLS regression of this variable. We will use filtering approach by Getis and Griffith (2002) based on the idea that it is possible to remove the spatial dependence from an autocorrelated variable and thus produce a new, spatially independent, variable. This filtered variable, log GDP per capita in our case, is then used in the log t test.

4 Results

4.1 Absolute convergence or convergence clubs?

In this work, we analyzed 275 European NUTS 2 regions across the new and old member states of the EU. We used the log t test by Phillips and Sul (2009), described above. The test and the associated clustering procedure were applied to the log GDP per capita (in PPS) between years 2000 and 2015. The data were obtained primarily from "Regional statistics by NUTS classification" database on the Eurostat website. As suggested by Phillips and Sul (2009), first 30% of the data was removed in order to give more weight to the latter part of the sample.

Applying the log t test, the convergence among all the regions of the EU is rejected with the log t coefficient equal to -20.57569. We can therefore see that there is no overall convergence in the EU in the measured period.

Then, we tested for the club convergence. Following the Phillips and Sul's clustering procedure, we initially obtained 10 convergence clubs and two diverging regions - Luxembourg and Inner London – West. As suggested by Bartkowska and Riedl (2012), we subsequently tried to merge the adjoining groups together and tested whether they converge. This process resulted in five convergence clubs with first two and the last convergence club left as they were but clubs 3 to 6 and 7 to 9 merged into two new clubs. Tables 1 and 3 show results of the procedure after and before merging.

As noted above for the log t test, the value of the coefficient b from equation 10, is interpreted as the speed of convergence and also shows sign and magnitude of the t-statistic, plays crucial role in the log t test. By analyzing the value of b for both versions of the test (5 and 10 clubs), we found out that the first convergence club shows strong relative (growth) convergence with significantly positive estimate of b equal to 0.26612. On the other hand, all other convergence clubs are rather weak. We can see negative, albeit insignificant, values of b coefficients suggesting relative convergence at very slow rate, as estimate of parameter α is not significantly different from 0

Table 1: Convergence club classification after merging

Club	N	log(t)	t value
Club 1	17	0.26612	2.693**
Club 2	36	-0.0148	-0.088
Club 3	93	-0.1114	-1.032
Club 4	98	-0.0625	-0.593
Club 5	29	-0.09851	-0.638

Note: **p<0.05

(Phillips and Sul, 2007). We cannot confirm level convergence within any of the clubs. The clubs before merging show higher value of t-statistics with all the b coefficients being positive. Nevertheless, the quantitative conclusions are the same for both versions - we cannot see any level convergence within the convergence clubs in the EU and we only observe stronger or weaker growth convergence.

We also plotted the convergence behaviour using relative transition coefficient h_{it} from 8, which traces out individual transition path of each cross-sectional unit with respect to a group average, for graphical investigation of relative convergence over time. Figures 3 and 4 show the relative transition paths for the first and the last convergence club. These pictures confirm that the convergence paths can be various, in spite of the ultimate convergence, with some regions rising from very low relative level, other descending from relatively higher level of GDP per capita to the group average. We can also note alternating states of convergence and divergence for some regions.

Figure 1 illustrates geographic composition of the convergence clubs in Europe. We can see that the capital regions tend to be part of the highest clubs. They are also very often surrounded by regions belonging to, sometimes much, lower clubs. This can be easily visible in case of the Paris region. There is also evident the previously suggested tendency of the clubs to cluster together, as can be demonstrated by the presence of continuous bright colored areas versus darker blue areas in the figure 1[2]. These results thus confirm previous findings about convergence clubs as made for example by Smetkowski and Wójcik (2012) or Bartkowska and Riedl (2012).

When trying to interpret the results, we can point out some significant

[2]Note, that there are also the outlying regions (Luxembourg and Inner London - West) included, forming the "Club 0"

differences within individual countries, expressed most often by the capital-periphery division and thus further confirming previous findings about significant regional disparities at country level. This is certainly also true for the CEE regions with some countries containing almost entire range of the convergence clubs (Poland, Romania). However, even Western Europe is not spared of this phenomenon (Italy or the UK are examples from the old countries).

We also found differences to the previous literature, for example, and in contrast to the conclusions of Bartkowska and Riedl (2012) about clubs clustering inside EU countries, we can see that most of the higher clubs without the capital cities seem to be very much geographically concentrated across the national borders. There is a clearly visible large area of prosperous regions from northern Italy to Benelux, with southern German regions in its heart. This area coincides strongly with so-called "Blue banana", a large urbanization corridor in Western Europe (Hospers, 2002). Worth noting is also the fact that the difference between the old and the new members in their club membership does not seem to be very strict, with large parts of rural France and Poland actually belonging to the same club.

At the same time, there are many CEE capitals in the two highest clubs, which leads us to conclusion that we do not see much evidence for the new-old division, as suggested in the previous research (Eckey and Türck, 2007). The final clubs' distribution seems more likely to speak for south-north division as the concentration of the lowest clubs is clearly the highest in the southern part of the EU. Performance of the capital cities and the emergence of the "Blue banana" also suggest an existence of broadly defined centre-periphery division.

5 Determinants of club membership

After finding the convergence clubs we tried to analyze factors that influence economic performance of the European regions. We used the previously found convergence clubs as dependent variable. Given the character of our dependent variable, which is discrete and can be logically ordered, we considered the ordered logistic regression framework to be the best for our analysis.

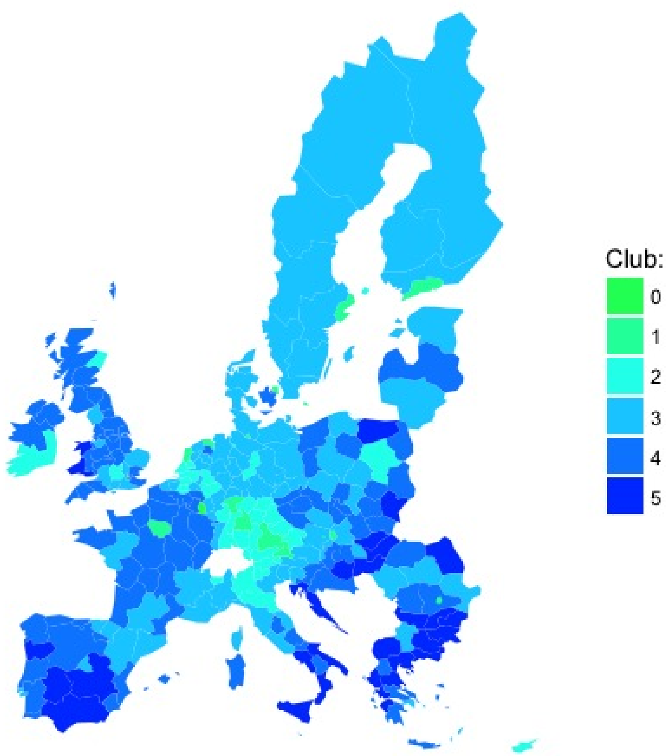

Figure 1: Convergence clubs and diverging regions (Club 0)

5.1 Explanatory variables

We used several explanatory variables for our analysis. We again worked with Eurostat data ("Regional statistics by NUTS classification"). Since we focus on the post-crisis development, and we want to investigate the role of sectoral composition, we used the data from 2008 to 2015 in the logistic regression instead of the 2000 - 2015 period used for the convergence analysis. This is mainly due to limited availability of these data for the new member states as well as changes in the Eurostat definition of Business services as discussed below. The choice of these variables follows New Economic Geography (NEG) and New Growth Theory (NGT), as well as previous research papers such as Mora (2008) or Bartkowska and Riedl (2012). Table 4 presents an overview of these variables with their names as they appear in the regression results below.

First of all, we included general population characteristics - population growth, use of which was inspired by neoclassical theory, and ratio between youth and elderly population. Furthermore, we used percentage of population with tertiary education and share of scientists in the active population as measures of regional population skills sophistication. We also included average number of patent applications to the European Patent Organization per million inhabitants in order to capture real innovation activity in the region. For all of the just mentioned variables we used their averages between years 2008 and 2015. Furthermore, we used number of regions within certain Euclidean distance d from a specific region, as set by the procedure of Getis and Ord's spatial filtering, variable aimed at expressing region's position with respect to the continent's geographical centre. Difference from EU average mean hourly earnings at NUTS 1 level for the year 2010 was used as possible measure of attractiveness for economic migration. In case of this variable, each NUTS 2 region was assigned the value of the respective NUTS 1 region.

Furthermore, we included variables measuring level of economic specialization in various sectors. The specialisation in agriculture and industry measured as share of population employed in these sectors over the overall employment for year 2008. Following Guastella and Timpano (2016) we defined specialization of an economy in Business Services as share of workers employed in BS in total employed population. This work use the definition of the Business services given by regulation of the European Parliament No. 298/2008. This regulation defines Business services as consisting of five divisions and three groups of NACE Rev. 2 classification (divisions 62, 69, 71, 73 and 78 and groups 58.2, 63.1 and 70.2). The data are not available in such detail in the Eurostat database for the Business services, therefore whole divisions (58, 63 and 70) are used as reasonable approximations for the respective groups. In the logistic regression, mentioned below, we work with a change in the specialization in Business services between years 2008 and 2014. The same time span was also preferred for other explanatory variables.

We also used percentage of individuals that have an interaction with public authorities via the internet. This should represent a general measure of quality of public services and openness as well as measure of how innovative a region is. Another tool for estimating the institutional quality was a dummy variable capturing presence of university between the top 150 world universities in 2008 in the region, we used the Shanghai university ranking (ARWU) for that year. As mentioned above, there is evidence for high level of economic inequality among the urban centres and other regions. We therefore included dummies for capital cities and regions with large urban areas, as they are defined by Eurostat. Last but not least, we measured whether there are headquarters of at least one of the 100 biggest EU companies as ranked by market capitalization by PwC the for the period between 2008 and 2014 (PwC, 2014).

Data for the new member states are available only since 2004 or 2007. Secondly, the Business services, are a composite of several categories of services and the data indexing of Eurostat changed between years 2007 and 2008 (from NACE Rev. 1 to Rev. 2 classification). This means, in practice, a change in the definition of BS itself. We therefore collected only the data since the year 2008.

Finally, due to missing observations, we have run the logistic regression on 245 out of 275 regions available with 30 regions removed due to high unavailability of the data. These were regions with either more than 6 observations in BS missing or those with low number of observations in all the other variables. Other missing variables were imputed [3]. It's worth noting that, even though these do not appear in the final logistic regression, we used state dummies for the data imputation, as we believe that these can help to take into account possible correlation among regions at the country level.

[3]We used several methods to interpolate the missing values. Apart from the geometric mean of employment in BS between years 2008 and 2014 as a measure of growth rate of Business services, we used k-Nearest Neighbours algorithm and factor analysis for data imputation.

5.2 Regression Results

Realizing that heteroscedasticity might be a problem in case of discrete choice models leading to not only incorrect standard errors but also to biased estimates, as shown by Williams (2009) and Keele and Park (2006), we included weights into the regression.These weights were supposed to model standard deviation of the dependent variable or precision with which we assign region to a certain convergence club. For this purpose, we used the relative transition coefficient h_{it} that expresses relative position of a region to the club's average. As the coefficient for all members of a club should converge to 1, we can reasonably expect that regions that keep closer to this value are less likely to end up in the club only due to a temporarily development, whereas regions that are fluctuating or keeping further away from the core of the club are less certain in their membership. Formally, we calculated the weights as

$$w_i = \frac{1}{\frac{1}{T}\sum_T |h_{it} - 1|}$$

Despite certain level of arbitrariness, we believe that this measure can reasonably capture the precision of the dependent variable, i.e. of assignment of particular region into its convergence club.

As a further robustness check, the model was estimated without those regions that had large number of missing observations, resulting in totally 245 regions estimated. We also included estimation that clustered the model's errors at the state level to achieve more trustworthy inference about the parameters (Cameron and Miller, 2015). Clustering the errors at the state level is quite intuitive solution even though certainly imperfect, as the method expects the errors between states to be uncorrelated, which is somehow against the expected results of economic integration in the EU at the regional level.

We estimated various specifications of our model (5 and 10 clubs variants for different imputation methods and weights). The regression results obtained for the various specifications show that the estimated coefficients and their significance somewhat differ, but in most cases we can reach similar qualitative conclusions. We therefore report the results for the clubs after merging, other versions can be found in the Appendix.

In table 2 we report the marginal effects our model for merged clubs, standard regression outputs are reported in the Appendix for brevity. Marginal effects are calculated at the means of the independent variables and show impact of one unit change in an explanatory variable on probability that certain region will enter particular club (Carroll, 2017).

Table 2: Marginal effects: Logistic regression with merged Clubs and FAMD[b] imputation

	Club 5	Club 4	Club 3	Club 2	Club 1
Population growth	-0.150***	-1.357***	1.000***	0.470***	0.037**
Population density	0.000**	0.000***	0.000***	0.000**	0.000
Young\old ratio	0.023	0.209*	-0.154	-0.072*	-0.006
Tertiary edu. share	-0.007	-0.059	0.044	0.021	0.002
Share of scientists	-4.206***	-38.13***	28.09***	13.21***	1.036**
Business Services	-0.569***	-5.160***	3.802***	1.787***	0.140**
Infrastructure	-0.075***	-0.678***	0.500**	0.235***	0.018**
Wage difference	0.000	-0.004	0.003	0.001	0.000
Internet use	0.000	-0.003	0.002	0.001	0.000
Old EU states	0.013	0.129	-0.081	-0.056	-0.005
Spec. agriculture	-0.084***	-0.760***	0.560***	0.263***	0.021**
Spec. industry	-0.082***	-0.747***	0.550***	0.259***	0.020**
Metrop. regions	0.009	0.094	-0.061	-0.039	-0.003
Capitals	-0.023**	-0.247***	0.093**	0.161**	0.015*
Top universities	0.034	0.217	-0.191	-0.055	-0.004
Corporations	-0.007	-0.066	0.045	0.026	0.002
Patent activity	0.000***	-0.002***	0.001***	0.001***	0.000**
Connectivity	-0.047	-0.425	0.313	0.147	0.012

Significance levels: *p<0.1; **p<0.05; **p<0.01

[a]Factorial Analysis for Mixed Data, an imputation method based on Principal Components Analysis

[b]Factorial Analysis for Mixed Data, an imputation method based on Principal Components Analysis

First thing we have noticed was that practically no explanatory variable seems to have strong impact on probability of entering the highest convergence club with all of them having the marginal effect for this club significantly lower than for the other clubs and most often very close to zero, if significant at all. The only possible exception can be the variable measuring share of scientists in the population but even in this case we see a significant decrease in the effect for the highest club. This conclusion is true also for the Business services, which diminish the probability of entering the lowest two clubs and have a positive impact on the clubs 3 and 2. However, their marginal effect for Club 1 is significantly lower. This may be explained by supportive role and nature of Business services. Hertog (2000) notes that, despite the Business Services provider plays a key role in facilitating the innovation process, the innovation itself is usually done by the original client. On the other hand, their contribution to innovation in other parts of the economy seems to be confirmed by our research, pointing to general symbiosis between Business services providers and their clients.

We can also see a strong effect of the variables connected with research activity such as the share of scientists and patent activity even though we cannot confirm strong influence of the dummy variable for the top universities. Its marginal effects also give rather counter-intuitive results with positive association with the lower clubs. One of the possible interpretations of these results could be an importance of applied research capable of producing applicable innovations from which benefits the regional business eco-system.

We can also spot expected and significantly positive impact of the capital cities. This seems to be supported by the results for the dummy indicating presence of large corporations that are often located in these cities. However, it is worth noting that the marginal effect of this dummy also diminishes for the very highest club and is not always significant across the model specifications. On the other hand, we found relatively weak association between the level of urbanization and the economic performance, as the population density has practically negligible impact and the dummy indicating presence of large metropolitan areas is often insignificant. Moreover, it seems to have positive association with lower clubs (this could capture the phenomena of transformation of former industrial areas). Interesting is also the positive impact of the population growth as given by the unweighted regression, suggesting large relocation of population from poorer to richer regions. Furthermore, we found no significant impact of the wage differences among the regions as well as the variable measuring share of population with tertiary education.

We do not find a robust evidence for the remaining variables. However, the results suggest that regions located near the hypothetical European geographic centre and with a good infrastructure are likely to end up in higher clubs. Also the internet usage variable intended to measure institutional quality and openness gives expected positive results, even though weak and mostly insignificant. Interesting is also the positive impact of specialization in agriculture.

6 Conclusion

This work aimed at deeper investigation of regional convergence in the European Union after the economic crisis of 2008. We applied a convergence test developed by Phillips and Sul (2009) which allows us to distinguish between several versions of convergence. It can recognize absolute as well as club convergence and also distinguishes income level convergence from relative (growth) convergence. The result of this analysis was then used as a dependent variable in logistic regression, employed for investigation of key determinants of regional economic performance. Our choice of variables was inspired by the New Economic Geography and New Growth Theory and extracted from Eurostat. We also included Business services - group of service activities ranging from legal consultancy to R&D services that was found in previous research to play an important role in innovation and knowledge diffusion.

The Phillips and Sul's test rejected the hypothesis of an overall convergence among the EU regions. We therefore continued our analysis searching for potential convergence clubs. Our final results gave us five distinct clubs converging in income growth rates instead of income level. Examining composition of these clubs, we found significant disparities in the club membership inside many countries, which suggests high economic inequality, clearly visible on the capital-periphery contrast, as the capital cities are typically converging to the highest clubs in both the old and the new countries of the EU. Furthermore, we found high concentration of regions belonging to the two highest clubs in relatively continuous area from northern Italy to the Benelux ("Blue Banana"). We consider both these findings as evidence for a centre-periphery division in the EU. We can also see a significant concentration of the regions from the lowest clubs in the Southern periphery of the EU.

We did not find substantial difference between the new and the old members regarding their convergence pattern. While the top clubs membership is limited to the capital cities in the CEE countries, similar phenomenon is visible among the old members as well. Also, as already mentioned, the lowest clubs are concentrated on the Southern rather than Eastern periphery. We therefore found the South-North division as more plausible than the Old-New one, while we consider the centre-periphery division the most prominent phenomenon.

The regression results show low impact of any of the estimated variables on membership in the highest club. This holds also for the innovation facilitating Business services, despite their overall positive effect on probability of entering higher clubs. We propose an explanation for this result. We also find a significant positive association of membership in the higher clubs with research and patent activities. Together with the results for the Business services, this possibly suggests focusing on applied research and innovation as a way to higher economic performance. Moreover, we confirmed an expected positive impact of capital cities but not of metropolitan areas in general, which supports the centre-periphery division.

Bringing our results together, we see further confirmation of the centre-periphery division in the EU. Mainly in its "capital versus rest" form, but coefficients of the connectivity variable, even though not always significant, also suggest validity of purely geographical form of this division (note, that precisely the regions inside the "Blue banana" fit the criteria of well built infrastructure as well as geographical centrality). Interpretation of the other results is far from straightforward, however, they give us also a hint of possible antidote for this dichotomy that seems to be overcome by regions that are able to gather substantial research activity and utilize this potential to create an innovation based economy. Such regions would definitely not be limited to the countries' capitals. The Business services with their role of innovation transmitters certainly fit into this pattern.

7 References

Luc Anselin. Spatial econometrics. In B. Baltagi, editor, *A Companion to Theoretical Econometrics*. Blackwell, Oxford, 2001.

Luc Anselin and Serge Rey. Properties of Tests for Spatial Dependence in Linear Regression Models. *Geographical analysis*, 23(2):112–131, 1991.

Nicholas Apergis, Ekaterini Panopoulou, and Chris Tsoumas. Old Wine in a New Bottle: Growth Convergence Dynamics in the EU. *Atlantic Economic Journal*, 38(2):169–181, 2010.

Monika Bartkowska and Aleksandra Riedl. Regional Convergence Clubs in Europe: Identification and Conditioning Factors. *Economic Modelling*, 29 (1):22–31, 2012.

Roberto Basile, Sergio De Nardis, and Alessandro Girardi. Regional Inequalities and Cohesion Policies in the European Union. *ISAE working papers*, 23, 2001.

Catherine Baumont, Cem Ertur, and Julie Le Gallo. Spatial Convergence Clubs and the European Regional Growth Process, 1980–1995. In Bernard Fingleton, editor, *European Regional Growth*, pages 131–158. Springer, 2003.

József Benedek and György Kocziszky. Paths of Convergence and Polarization in the Visegrád Countries. In Thilo Lang, Sebastian Henn, Kornelia Ehrlich, and Wladimir Sgibnev, editors, *Understanding Geographies of Polarization and Peripheralization*, pages 217–234. Springer, 2015.

Mihály Tamás Borsi and Norbert Metiu. The Evolution of Economic Convergence in the European Union. *Empirical Economics*, 48(2):657–681, 2015.

A Colin Cameron and Douglas L Miller. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372, 2015.

Nathan Carroll. *oglmx*: A Package for Estimation of Ordered Generalized Linear Models. *CRAN, The Comprehensive R Archive Network*, 2017.

Sheila A Chapman, Stefania Cosci, and Loredana Mirra. Income Dynamics in an Enlarged Europe: the Role of Capital Regions. *The Annals of Regional Science*, 48(3):663–693, 2012.

Nicoletta Corrocher and Lucia Cusmano. The 'KIBS engine' of Regional Innovation Systems: Empirical Evidence from European Regions. *Regional Studies*, 48(7):1212–1226, 2014.

Jesús Crespo Cuaresma, Gernot Doppelhofer, and Martin Feldkircher. The Determinants of Economic Growth in European Regions. *Regional Studies*, 48(1):44–67, 2014.

Eleonora Cutrini. Economic integration, structural change, and uneven development in the european union. *Structural Change and Economic Dynamics*, 50:102–113, 2019.

Sandy Dall'Erba and Julie Le Gallo. Regional Convergence and the Impact of European Structural Funds over 1989–1999: A Spatial Econometric Analysis. *Papers in Regional Science*, 87(2):219–244, 2008.

Benoît Desmarchelier, Faridah Djellal, and Faïz Gallouj. Knowledge Intensive Business Services and Long Term Growth. *Structural Change and Economic Dynamics*, 25:188–205, 2013.

Hans Friedrich Eckey and Matthias Türck. Convergence of EU-Regions. A Literature Report/Convergencia de las regiones de la UE. Una revisión de la literatura. *Investigaciones regionales*, (10):5–32, 2007.

Manfred M Fischer and James P LeSage. A Bayesian Space-Time Approach to Identifying and Interpreting Regional Convergence Clubs in Europe. *Papers in Regional Science*, 94(4):677–702, 2015.

Ulrich Fritsche and Vladimir Kuzin. Analysing Convergence in Europe Using the Non-Linear Single Factor Model. *Empirical Economics*, 41(2):343–369, 2011.

Oded Galor. Convergence? Inferences from Theoretical Models. *The Economic Journal*, 106(437):1056–1069, 1996.

Kurt Geppert and Andreas Stephan. Regional Disparities in the European Union: Convergence and Agglomeration. *Papers in Regional Science*, 87 (2):193–217, 2008.

Arthur Getis and Daniel A Griffith. Comparative Spatial Filtering in Regression Analysis. *Geographical analysis*, 34(2):130–140, 2002.

Arthur Getis and J Keith Ord. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical analysis*, 24(3):189–206, 1992.

Madhusudan Ghosh, Atanu Ghoshray, and Issam Malki. Regional Divergence and Club Convergence in India. *Economic Modelling*, 30:733–742, 2013.

Gianni Guastella and Francesco Timpano. Knowledge, Innovation, Agglomeration and Regional Convergence in the EU: Motivating Place-Based Regional Intervention. *Review of Regional Research*, 2(36):121–143, 2016.

Pim den Hertog. Knowledge-Intensive Business Services as co-Producers of Innovation. *International journal of innovation management*, 4(04):491–528, 2000.

Gert-Jan Hospers. Beyond the Blue Banana? Structural Change in Europe's Geo-Economy. In *ERSA Conference Papers*. European Regional Science Association (ERSA), 2002.

Simona Iammarino, Andrés Rodríguez-Pose, and Michael Storper. Regional inequality in europe: evidence, theory and policy implications. *Journal of economic geography*, 19(2):273–298, 2019.

Luke Keele and David K Park. Difficult Choices: an Evaluation of Heterogeneous Choice Models. In *APSA 2004 Annual Meeting Paper*. American Political Science Association 2004 Annual Meeting, 2006.

Henk Kox and Luis Rubalcaba. The Contribution of Business Services to European Economic Growth. In Henk Kox and Luis Rubalcaba, editors, *Business Services in European Economic Growth*, pages 74–94. Springer, 2007.

Esfandiar Maasoumi and Le Wang. Economic Reform, Growth and Convergence in China. *The Econometrics Journal*, 11(1):128–154, 2008.

Stefano Magrini. Regional (di) Convergence. *Handbook of regional and urban economics*, 4:2741–2796, 2004.

Vassilis Monastiriotis. Regional Growth Dynamics in Central and Eastern Europe. *LEQS Paper*, (33), 2011.

Mercedes Monfort, Juan Carlos Cuestas, and Javier Ordonez. Real Convergence in Europe: A Cluster Analysis. *Economic Modelling*, 33:689–694, 2013.

Toni Mora. Factors Conditioning the Formation of European Regional Convergence Clubs. *The Annals of Regional Science*, 42(4):911–927, 2008.

Emmanuel Muller and Andrea Zenker. Business Services as Actors of Knowledge Transformation: The Role of KIBS in Regional and National Innovation Systems. *Research policy*, 30(9):1501–1516, 2001.

Peter CB Phillips and Donggyu Sul. Transition Modeling and Econometric Convergence Tests. *Econometrica*, 75(6):1771–1855, 2007.

Peter CB Phillips and Donggyu Sul. Economic Transition and Growth. *Journal of Applied Econometrics*, 24(7):1153–1185, 2009.

Carlos Pinho, Carlos Andrade, and Maria de Fátima Pinho. Regional Growth Transitions and the Evolution of Income Disparities in Europe. *Urban Public Economics Review*, (13):66, 2010.

PwC. Top 100 companies Europe and Germany, 2014. Retrieved on from `https://www.pwc.de/de/kapitalmarktorientierte-unternehmen/assets/healthcare-und-` on 17/11/2017.

Korneliusz Pylak and Dariusz Majerek. Tracing Regional Economic Evolutions. KIS Impact on Growth. *Procedia Economics and Finance*, 14: 515–524, 2014.

Domingo Rodríguez-Benavides, Francisco López-Herrera, and Francisco Venegas-Martínez. Are there Economic Convergence Clubs in Latin America? *Journal of Economics*, 2(3):113–123, 2014.

Xavier X Sala-i Martin. Regional Cohesion: Evidence and Theories of Regional Growth and Convergence. *European Economic Review*, 40(6):1325–1352, 1996.

Maciej Smętkowski. Regional Disparities in Central and Eastern European Countries: Trends, Drivers and Prospects. *Europe-Asia Studies*, 65(8): 1529–1554, 2013.

Maciej Smetkowski and Piotr Wójcik. Regional Convergence in Central and Eastern European countries: A Multidimensional Approach. *European Planning Studies*, 20(6):923–939, 2012.

Dóra Szendi. The Convergence Process and the Effects of the Economic Crisis in Central-Eastern Europe. *Romanian Review of Regional Studies*, 9(1), 2013.

Konrad Von Lyncker and Rasmus Thoennessen. Regional club convergence in the eu: evidence from a panel data analysis. *Empirical Economics*, 52 (2):525–553, 2017.

Richard Williams. Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients across Groups. *Sociological Methods & Research*, 37(4):531–559, 2009.

Grigoris Zarotiadis and Aristea Gkagka. European Union: a Diverging Union? *Journal of Post Keynesian Economics*, 35(4):537–568, 2013.

8 Appendix A: Composition of the Convergence Clubs

- Club 1:
  Wien, Salzburg, Région de Bruxelles-Capitale / Brussels Hoofdstedelijk Gewest, Praha, Stuttgart, Oberbayern, Hamburg, Darmstadt, Hovedstaden, Helsinki-Uusimaa, Île de France, Groningen, Noord-Holland, Bucuresti - Ilfov, Stockholm, Bratislavský kraj, Inner London - East

- Club 2:
  Oberösterreich, Tirol, Vorarlberg, Prov. Antwerpen & Vlaams-Brabant, Prov. Brabant Wallon, Kypros, Karlsruhe, Freiburg, Tübingen, Niederbayern, Oberpfalz, Oberfranken, Mittelfranken, Unterfranken, Schwaben, Berlin, Bremen, Braunschweig, Düsseldorf, Köln, Rheinhessen-Pfalz, Southern and Eastern, Valle d'Aosta/Vallée d'Aoste, Lombardia, Provincia Autonoma di Bolzano/Bozen, Provincia Autonoma di Trento, Emilia-Romagna, Utrecht, Zuid-Holland, Noord-Brabant, Mazowieckie, Berkshire, Buckinghamshire and Oxfordshire, North Eastern Scotland

- Club 3:
  Burgenland (AT), Niederösterreich, Kärnten, Steiermark, Prov. Limburg (BE), Prov. Oost-Vlaanderen, Prov. West-Vlaanderen, Yugozapaden, Jihovýchod, Brandenburg, Gießen, Kassel, Mecklenburg-Vorpommern, Hannover, Lüneburg, Weser-Ems, Münster, Detmold, Arnsberg, Koblenz, Trier, Saarland, Dresden, Chemnitz, Leipzig, Sachsen-Anhalt, Schleswig-Holstein, Thüringen, Syddanmark, Midtjylland, Eesti, Attiki, Notio Aigaio, País Vasco, Comunidad Foral de Navarra, Aragón, Comunidad de Madrid, Cataluña, Canarias (ES), Länsi-Suomi, Etelä-Suomi, Alsace, Pays de la Loire, Midi-Pyrénées, Rhône-Alpes, Provence-Alpes-Côte d'Azur, Corse, Közép-Magyarország, Piemonte, Liguria, Veneto, Friuli-Venezia Giulia, Toscana, Marche, Lazio, Lietuva, Friesland (NL), Overijssel, Gelderland, Flevoland, Zeeland, Limburg (NL), Lódzkie, Slaskie, Wielkopolskie, Dolnoslaskie, Pomorskie, Área Metropolitana de Lisboa, Centru, Sud-Est, Vest, Östra Mellansverige, Småland med öarna, Sydsverige, Västsverige, Norra Mellansverige, Mellersta Norrland, Övre Norrland, Zahodna Slovenija, Západné Slovensko, Cumbria, Cheshire (NUTS 2006), East Anglia, Bedfordshire and Hertfordshire, Outer London - West and North West, Surrey, East and West

Sussex, Hampshire and Isle of Wight, Gloucestershire, Wiltshire and Bristol/Bath area

- Club 4:
Prov. Liège, Střední Čechy, Jihozápad, Severozápad, Severovýchod, Střední Morava, Moravskoslezsko, Voreio Aigaio, Kriti, Galicia, Principado de Asturias, Cantabria, La Rioja, Castilla y León, Comunidad Valenciana, Illes Balears, Región de Murcia, Champagne-Ardenne, Picardie, Haute-Normandie, Centre (FR), Basse-Normandie, Bourgogne, Lorraine, Franche-Comté, Bretagne, Poitou-Charentes, Aquitaine, Limousin, Auvergne, Languedoc-Roussillon, Közép-Dunántúl, Nyugat-Dunántúl, Border, Midland and Western, Abruzzo, Molise, Basilicata, Sardegna, Umbria, Latvija, Drenthe, Malopolskie, Lubelskie, Swietokrzyskie, Podlaskie, Zachodniopomorskie, Lubuskie, Opolskie, Kujawsko-Pomorskie, Centro (PT), Nord-Vest, Sud - Muntenia, Sud-Vest Oltenia, Vzhodna Slovenija, Stredné Slovensko, Východné Slovensko, Tees Valley and Durham, Northumberland and Tyne and Wear, Greater Manchester, Lancashire, Merseyside, East Yorkshire and Northern Lincolnshire, North Yorkshire, South Yorkshire, West Yorkshire, Derbyshire and Nottinghamshire, Leicestershire, Rutland and Northamptonshire, Lincolnshire, Herefordshire, Worcestershire and Warwickshire, Shropshire and Staffordshire, West Midlands, Essex, Outer London - East and North East, Outer London - South, Kent, Dorset and Somerset, Cornwall and Isles of Scilly, Devon, East Wales, Eastern Scotland, South Western Scotland, Highlands and Islands, Northern Ireland (UK)

- Club 5:
Severozapaden, Severen tsentralen, Severoiztochen, Yugoiztochen, Yuzhen tsentralen, Kentriki Makedonia, Castilla-la Mancha, Extremadura, Andalucía, Dél-Dunántúl, Észak-Magyarország, Észak-Alföld, Dél-Alföld, Campania, Puglia, Calabria, Sicilia, Podkarpackie, Warminsko-Mazurskie, Norte, Nord-Est, West Wales and The Valleys

9 Appendix B: Tables and Regression results

9.1 Relative Transition Paths

Table 3: Convergence club classification before merging

Club	N	log(t)	t value
Club 1	17	0.26612	2.693**
Club 2	36	-0.0148	-0.088
Club 3	25	0.3725	2.334**
Club 4	24	0.1924	0.974
Club 5	30	0.02397	0.223
Club 6	14	0.04248	0.328
Club 7	28	0.08075	0.720
Club 8	30	0.2519	1.895**
Club 9	40	0.1067	0.692
Club 10	29	-0.09851	-0.638

Note: **p<0.05

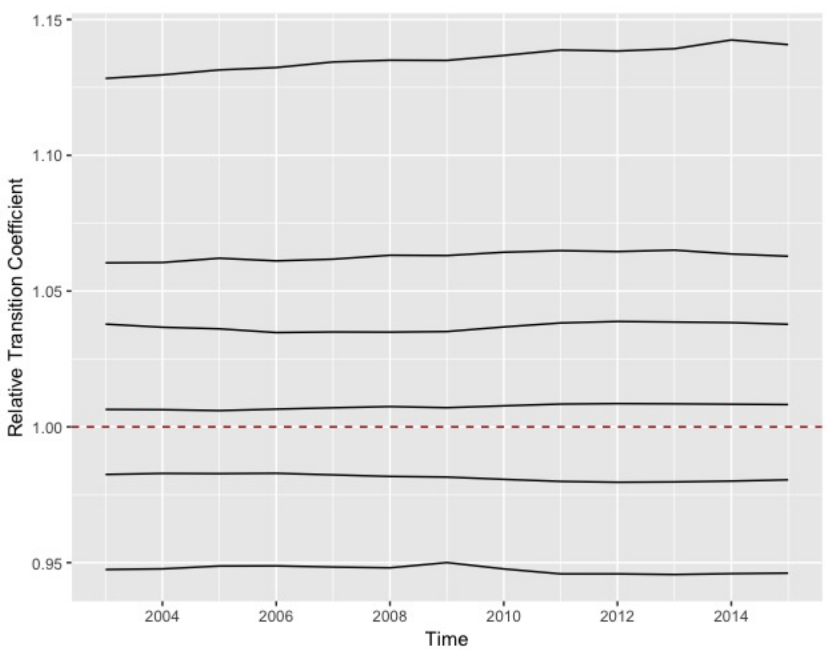

Figure 2: Average transition paths across merged convergence clubs. Diverging regions are represented by the highest transition curve.

Table 4: Explanatory variables - description

Regression variable	Description
Population growth	Average grow rate of total population
Population density	Average population density
Young\old ratio	Ratio between young (under 15) and elderly (over 65)
Tertiary edu. share	Average percentage of population with tertiary education
Share of scientists	Share of scientists in active population
Business Services	Change in specialization in Business Services (2008 - 2014)
Infrasturcture	Number of all vehicles per capita (in 2008)
Wage difference	Difference from the EU average of mean hourly earnings (at NUTS 1 level, in 2010)
Internet use	Average percentage of individuals that had an interaction with public authorities via the internet within last year (in 2013 - 2016).
Old EU states	Dummy indicating regions from the old EU members
Spec. agriculture	Share of population from 15 to 64 years employed in agriculture in all employed (in 2008)
Spec. Industry	Share of population from 15 to 64 years employed in Industry in all employed (in 2008)
Metrop. regions	Dummy indicating presence of a large urban area
Capitals	Dummy indicating regions with capital cities
Top universities	Dummy indicating presence of a university between the top 150 world universities in 2008
Corporations	Dummy indicating presence of a headquarter of at least one of the 100 biggest EU companies
Patent activity	Average number of patents per million inhabitants
Connectivity	Number of regions within Euclidean distance from certain region as set by the procedure of Getis and Ord

Note: Unless stated otherwise, the variables are used as averages between 2008-2015.

Table 5: Marginal effects: Weighted Logistic Regression with merged Clubs and kNN[b] imputation

	Club 5	Club 4	Club 3	Club 2	Club 1
Population growth	-0.001	-0.020	0.012	0.009	0.000
Population density	0.000*	-0.000***	0.000***	0.000***	0.000
Young\old ratio	0.002	0.082	-0.048	-0.034	-0.001
Tertiary edu. share	-0.001	-0.039	0.023	0.017	0.000
Share of scientists	-1.629*	-58.53***	34.79***	24.95***	0.418
Business Services	-0.108	-3.862	2.296	1.646**	0.028
Infrasturcture	-0.015	-0.538	0.320	0.229	0.004
Wage difference	0.000	-0.002	0.001	0.000	0.000
Internet use	0.000	-0.005*	0.003**	0.002	0.000
old EU states	-0.002	-0.071	0.045	0.027	0.000
Spec. agriculture	-0.059*	-2.135**	1.269***	0.910*	0.015
Spec. Industry	-0.017	-0.610	0.362	0.260	0.004
Metrop. regions	0.003	0.123*	-0.056**	-0.070	-0.001
Capitals	-0.006*	-0.254***	0.036***	0.220**	0.005
Top universities	0.008	0.210	-0.153	-0.064*	-0.001
Corporations	-0.005*	-0.210***	0.049	0.162*	0.003
Patent activity	0.000	-0.002***	0.000***	0.000	0.000
Connectivity	-0.041**	-1.504***	0.894	0.641**	0.011

Significance levels: *p<0.1; **p<0.05; ***p<0.05

[a]k-Nearest Neighbours
[b]k-Nearest Neighbours

Table 6: Regression results: Logistic regression with merged Convergence Clubs and FAMD imputation

Dependent variable:	Conververgence Club		
Population growth	6.805***	Old EU states	−0.691
	(0.949)		(0.544)
Population density	0.001**	Spec. agriculture	3.810***
	(0.0003)		(0.400)
Young\old ratio	−1.049*	Spec. industry	3.744***
	(0.630)		(0.203)
Tertiary edu. share	0.298	Metrop. regions	−0.494
	(0.838)		(0.432)
Share of scientist	191.189***	Capitals	1.535***
	(0.017)		(0.569)
Business Services	25.876***	Top universities	−1.053
	(0.025)		(0.670)
Infrastructure	3.400***	Corporations	0.345
	(0.666)		(0.390)
Wage difference	0.022	Patent activity	0.010***
	(0.042)		(0.002)
Internet use	0.013	Connectivity	2.133
	(0.024)		(1.362)
Observations	244		

Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table 7: Regression results: Weighted Logistic Regression with merged Convergence Clubs and kNN imputation

Dependent variable:	Conververgence Club		
Population growth	0.101 (6.374)	Old EU states	0.346 (1.383)
Population density	0.0008*** (0.0002)	Spec. agriculture	10.73** (4.971)
Young\old ratio	−0.4111 (1.186)	Spec. industury	−3.066 (4.319)
Tertiary edu. share	0.200 (3.146)	Metrop. regions	−0.687 (0.422)
Share of scientists	294.3*** (77.55)	Capitals	1.692** (0.673)
Business Services	19.42 (12.59)	Top universities	−0.963 (0.643)
Infrastructure	2.70 (2.196)	Corporations	1.329*** (0.514)
Wage difference	0.009 (0.057)	Patent activity	0.010*** (0.002)
Internet use	0.029 (0.018)	Connectivity	7.562*** (2.039)
Observations	219		

Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table 8: Regression results: Logistic regression with Convergence Clubs before merging and FAMD imputation

Dependent variable:	Conververgence Club		
Population growth	5.790***	Old EU states	−1.257**
	(0.859)		(0.545)
Population density	0.001***	Spec. agriculture	2.230***
	(0.0003))		(0.324)
Young\old ratio	−1.494***	Spec. industry	4.650***
	(0.511)		(0.179)
Tertiary edu. share	0.602	Metrop. regions	−0.537
	(0.854)		(0.327)
Share of scientists	181.929***	Capitals	1.583***
	(0.017)		(0.581)
Business Services	27.782***	Top universities	−1.082*
	(0.020)		(0.409)
Infrastructure	3.992***	Corporations	0.339
	(0.455)		(0.414)
Wage difference	0.055	Patent activity	0.010***
	(0.039)		(0.003)
Internet use	0.012	Connectivity	1.464
	(0.024)		(1.347)
Observations:	244		

Significance levels: *p<0.1; **p<0.05; ***p<0.01

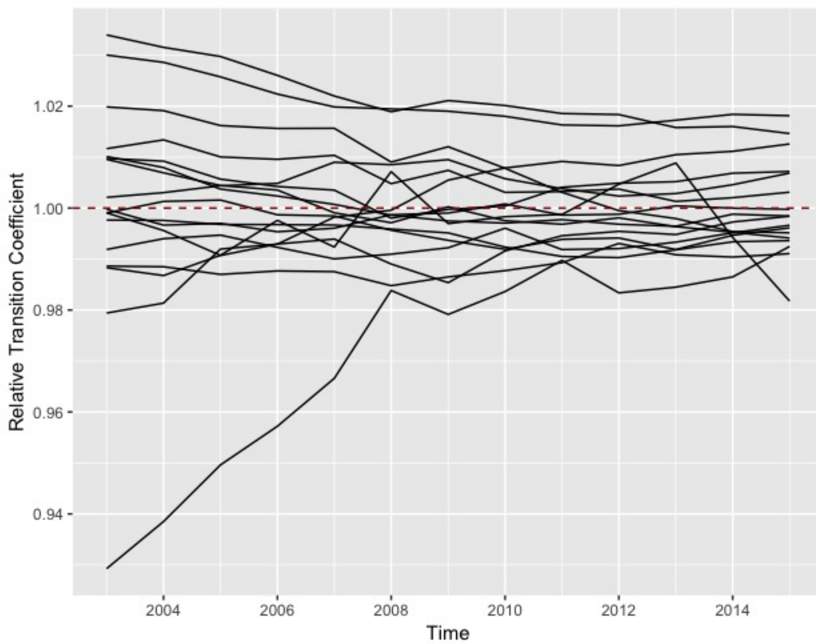

Figure 3: Transition paths for members of the first Club

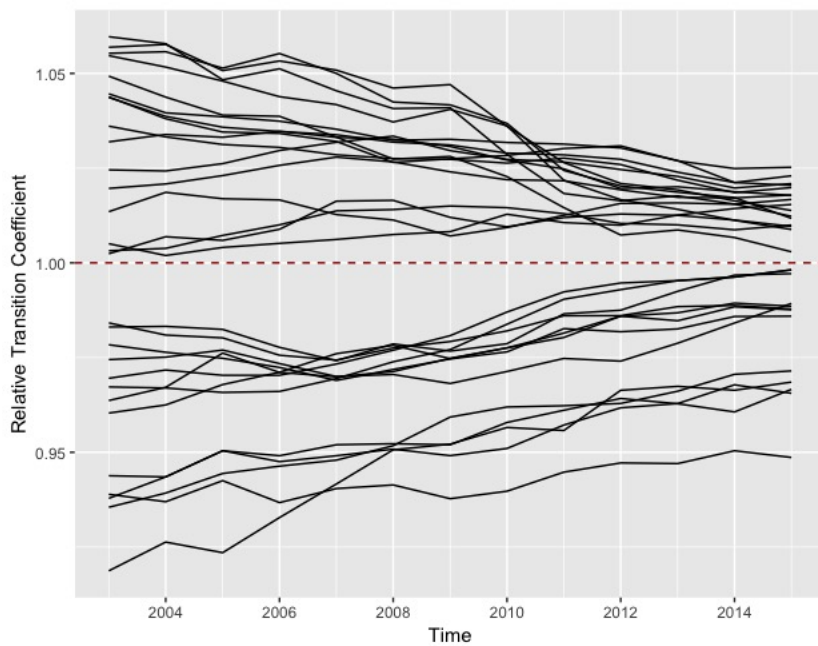

Figure 4: Transition paths for members of the fifth Club

IES Working Paper Series

2021

1. Mahir Suleymanov: *Foreign Direct Investment in Emerging Markets: Evidence from Russia since the 2000s*
2. Lenka Nechvátalová: *Multi-Horizon Equity Returns Predictability via Machine Learning*
3. Milan Scasny, Matej Opatrny: *Elasticity of Marginal Utility of Consumption: The Equal-Sacrifice Approach Applied for the Czech Republic*
4. Javier Garcia-Bernardo, Petr Jansky and Vojtech Misak: *Common Agricultural Policy Beneficiaries: Evidence of Inequality from a New Data Set*
5. Petr Jakubik, Saida Teleu: *Suspension of Insurers' Dividends as a Response to the Covid-19 Crisis: Evidence from Equity Market*
6. Boris Fisera, Menbere Workie Tiruneh, David Hojdan: *Currency Depreciations in Emerging Economies: A Blessing or a Curse for External Debt Management?*
7. Vojtech Molnar: *Price Level Targeting with Imperfect Rationality: A Heuristic Approach*
8. Alex Cobham, Tommaso Faccio, Javier Garcia-Bernardo, Petr Jansky, Jeffery Kadet, Sol Picciotto: *A Practical Proposal to End Corporate Tax Abuse: METR, a Minimum Effective Tax Rate for Multinationals*
9. Evžen Kočenda, Ichiro Iwasaki: *Bank Survival Around the World: A Meta-Analytic Review*
10. Michal Kuchta: *Scenario Generation for IFRS9 Purposes using a Bayesian MS-VAR Model*
11. Jozef Barunik, Josef Kurka: *Frequency-Dependent Higher Moment Risks*
12. Petr Pleticha: *Who Benefits from Global Value Chain Participation? Does Functional Specialization Matter?*
13. Alex Cobham, Petr Jansky, Chris Jones, Yama Temouri: *An Evaluation of the Effects of the European Commission's Proposals for the Common Consolidated Corporate Tax Base*
14. Javier Garcia-Bernardo, Petr Jansky: *Profit Shifting of Multinational Corporations Worldwide*
15. Tomas Domonkos, Boris Fisera and Maria Siranova: *Income Inequality as Long-term Conditioning Factor of Monetary Transmission to Bank Interest Rates in EA Countries*
16. Petr Jansky, Natalia Li: *Improving the Corruption Perceptions Index: Additional Data Sources and Their Effects*
17. Dominika Ehrenbergerova, Josef Bajzik, Tomas Havranek: *When Does Monetary Policy Sway House Prices? A Meta-Analysis*
18. Daniel Kolář: *Inequality in Pre-Industrial Urban Bohemia: The City of Budweis*
19. Barbora Malinska: *Forecasting Sovereign Bond Realized Volatility Using Time-Varying Coefficients Model*

20. Jan Pintera: *Regional Convergence in the European Union: What are the Factors of Growth?*

All papers can be downloaded at: http://ies.fsv.cuni.cz.

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz http://ies.fsv.cuni.cz