

Charles University
Faculty of Social Sciences
Institute of Economic Studies



MASTER'S THESIS

**Regional Convergence in the European
Union: Do the Business Services Make
the Difference?**

Author: **Bc. Jan Pintera**

Supervisor: **PhDr. Jaromír Baxa, Ph.D.**

Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, December 28, 2017

Signature

Acknowledgments

I would like to kindly thank to Jaromír Baxa for advice and guidance on this thesis. I also express my gratitude to Tomáš Havránek for useful remarks on the econometric part of the work.

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 thesis provides novel evidence on convergence in the EU while focusing on development at regional level after the Great Recession. It uses recently developed log t convergence test by Phillips & Sul (2007). Our findings speak against the convergence in level of income per capita among the European regions and give us five clubs of regions converging in their income growth rates instead. 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. The main explanatory variable of interest were Business Services (BS), a dynamic sector of the economy with presumably strong positive effect on regional innovative potential. We found positive effect of BS on membership in higher convergence clubs. Yet, this effect seems to diminish for the very highest club.

JEL Classification C23, C40, R11

Keywords Club Convergence, European Regions, log t test, Logistic regression

Author's e-mail jan.pintera@gmail.com

Supervisor's e-mail jaromir.baxa@fsv.cuni.cz

Abstrakt

I přes dlouhodobě se prohlubující hospodářskou spolupráci mezi členskými zeměmi a regiony zůstávají závěry empirického výzkumu ohledně konvergence nejasné. Tato práce si klade za cíl doplnit stávající výzkum o nové poznatky pomocí nedávno vyvinuté metody - $\log t$ konvergenčního testu (Phillips & Sul 2007) s využitím dat z období po ekonomické krizi z roku 2008. Dosažené výsledky práce nepotvrdily konvergenci v hladině HDP na obyvatele, zato však bylo nalezeno pět skupin (klubů) tvořených regiony konvergujícími v míře hospodářského růstu. Geografické rozložení těchto klubů potvrdilo předpokládanou vysokou míru regionální nerovnosti v jednotlivých zemích a zároveň ukázalo na existenci většího seskupení bohatých regionů v metropolitní oblasti západní Evropy. Práce se rovněž věnuje analýze faktorů podmiňujících členství v jednotlivých konvergenčních klubech pomocí logistické regrese. Jako hlavní vysvětlující proměnná byly použity tzv. Business services (BS), dynamický sektor služeb, v literatuře často spojovaný se silným inovačním potenciálem. V jejich případě byl nalezen pozitivní efekt na členství ve vyšších klubech, který měl ovšem tendenci slábnout v případě klubu nejvyššího.

Klasifikace JEL

C23, C40, R11

Klíčová slova

Regionalní konvergence, Evropská Unie,
 $\log t$ test, Logistická regrese

E-mail autora

jan.pintera@gmail.com

E-mail vedoucího práce

jaromir.baxa@fsv.cuni.cz

Contents

| | |
|--|-----------|
| List of Tables | viii |
| List of Figures | ix |
| Acronyms | x |
| Thesis Proposal | xi |
| 1 Introduction | 1 |
| 2 Literature review | 5 |
| 2.1 The convergence hypothesis and modelling frameworks for its empirical measurement | 5 |
| 2.2 Empirical analyses of convergence among EU regions | 7 |
| 2.3 Stylised facts about convergence in CEE regions | 9 |
| 2.4 Business services - introduction and definition | 11 |
| 2.5 Business services as a factor of economic development | 14 |
| 3 Methodology | 16 |
| 3.1 The log t test by Phillips and Sul - empirical properties | 16 |
| 3.2 The log t test - modelling framework | 17 |
| 3.3 The log t test - formation of convergence clubs | 21 |
| 3.4 Spatial correlation - issues and solutions | 23 |
| 3.5 Impact of the Business services - methodology | 24 |
| 4 Data and Variables | 26 |
| 4.1 Explanatory variables | 26 |
| 4.2 Treatment of missing observations | 28 |
| 5 Results | 33 |
| 5.1 Absolute convergence or convergence clubs? | 33 |

| | | |
|----------|--|-----------|
| 5.2 | Determinants of club membership - estimation framework | 40 |
| 5.3 | Drivers of Convergence - discussion | 43 |
| 6 | Conclusion | 46 |
| | Bibliography | 53 |
| A | Composition of the Convergence Clubs | I |
| B | Relative Transition Paths | IV |
| C | Regression Outputs | X |

List of Tables

| | | |
|-----|--|------|
| 4.1 | Explanatory variables - description | 29 |
| 5.1 | Convergence club classification after merging | 34 |
| 5.2 | Convergence club classification before merging | 37 |
| 5.3 | Marginal effects: Logistic regression with merged Clubs and FAMD imputation | 42 |
| 5.4 | Marginal effects: Weighted Logistic Regression with merged Clubs and kNN imputation | 44 |
| C.1 | Regression results: Logistic regression with Convergence Clubs before merging and FAMD imputation | X |
| C.2 | Marginal effects: Logistic regression with Convergence Clubs before merging and FAMD imputation | XI |
| C.3 | Regression results: Logistic Regression with merged Conver- gence Clubs and kNN imputation | XII |
| C.4 | Marginal Effects: Logistic Regression with merged Convergence Clubs and kNN imputation | XIII |
| C.5 | Regression results: Weighted Logistic Regression with merged Convergence Clubs and kNN imputation | XIV |
| C.6 | Regression results: Logistic regression with merged Convergence Clubs and FAMD imputation | XV |

List of Figures

| | | |
|------|--|------|
| 5.1 | Average transition paths across merged convergence clubs | 35 |
| 5.2 | Transition paths for members of the first Club | 35 |
| 5.3 | Transition paths for members of the fifth Club | 36 |
| 5.4 | Convergence clubs and diverging regions (Club 0) | 39 |
| | | |
| B.1 | Transition paths for members of Club 2 (merged) | IV |
| B.2 | Transition paths for members of Club 3 (merged) | V |
| B.3 | Transition paths for members of Club 4 (merged) | V |
| B.4 | Transition paths for members of Club 3 before merging | VI |
| B.5 | Transition paths for members of Club 4 before merging | VI |
| B.6 | Transition paths for members of Club 5 before merging | VII |
| B.7 | Transition paths for members of Club 6 before merging | VII |
| B.8 | Transition paths for members of Club 7 before merging | VIII |
| B.9 | Transition paths for members of Club 8 before merging | VIII |
| B.10 | Transition paths for members of Club 9 before merging | IX |

Acronyms

BS Business services

CEEC Central and Eastern European Countries

FAMD Factor Analysis for Mixed Data

KIBS Knowledge Intensive Business Services

KIS Knowledge Intensive Services

kNN k-Nearest Neighbours

NUTS Nomenclature of Territorial Units for Statistics

NEG New Economic Geography

NGT New Economic Theory

Master's Thesis Proposal

| | |
|-----------------------|--|
| Author | Bc. Jan Pintera |
| Supervisor | PhDr. Jaromír Baxa, Ph.D. |
| Proposed topic | Regional convergence in the Central and Eastern European members of the EU |

Motivation Economic convergence was undeniably one of the most important reasons for entering the EU for most of its new members from so-called CEE countries. On the other hand, even more than ten years after the new countries became members, we can still see a gap between them and the older EU countries. Moreover, if we simply look at the GDP growth figures, we do not have to see convergence at all in some periods. Question of actual convergence therefore does not seem straightforward and remains to be of the highest economic as well as political interest.

There has been many works trying to estimate and verify existence of convergence using various approaches. The main source of methodology for this work will be Phillips and Sul's (2007) framework which was since then adopted by many empirical research papers to estimate the convergence clubs on the regional as well as country level. Example of the latter can be Borsi and Metiu (2013) whose results are most importantly suggesting that CEE countries do not experience growth sufficient to catch up with the old EU members. Regional convergence is then examined by Bartkowska and Riedl (2012) who note several interesting findings about regional convergence – namely tendency of converging regions to be spatially clustered or finding that metropolitan areas tend to belong to clusters with higher GDP. Among other aims of this thesis will therefore be to verify these findings while including post-crisis data to the analysis.

Another aspect of economic convergence is that whether a country is converging or not can be changing over time and especially the state after the 2008 crisis is of great interest. This work aims to combine the techniques for finding and describing the convergence clubs with theory describing sectoral specialization of different regions and thus answer the question whether the convergence exists as well as whether it is driven by long term structural factors or whether much more volatile business

cycle factors play a role as well. To achieve this goal, work of Meliciani and Savona (2014) will be used. The authors investigate NUTS2 level concentration of so-called business services. Concentration on this type of services should be strongly growth promoting. This work will therefore attempt to find the link between different levels of regional specialization in business services and presence the respective regions in different clusters.

Hypotheses

Hypothesis #1: The CEEC NUTS2 regions converge to the EU 15 average.

Hypothesis #2: The main agglomerations areas form a distinct convergence club.

Hypothesis #3: The regions with higher Business service concentration converge to higher equilibria than old manufacturing or agriculture areas. decreases in time.

Hypothesis #4: Areas with similar BS concentration converge to the same equilibrium in the old and new EU countries

Methodology The aim of this diploma thesis is to examine regional economic convergence of the per capita income in the CEE countries. The main goals of this thesis are to verify existence of regional converge between the old and new EU countries on NUTS 2 level and to test the hypothesis that the regional convergence clubs are created along different levels of specialization in so called business services (BS). The test used to detect the convergence clusters is based mainly on a variation of the neoclassical growth model and was described by Phillips and Sul (2007). This thesis will apply their methodology to the CEE countries in order to test multiple hypotheses about income convergence. Subsequently, some of the alternative methods for finding the convergence clubs (Markov chains, Kernel analysis, Clustering and Bayesian methods) will be used in order to compare the club composition and thus check robustness of the results. The primary result will be the convergence clubs themselves – which allows us to test where and which regions are converging to the EU or EU 15 average. Secondly the thesis aims to test whether we can find similar convergence dynamics within the CEE regions as was found by previous papers for the western European countries. Furthermore, the thesis will attempt to verify a connection between the economic convergence and specialization in one of the most productive sector of services – the business services as described by Meliciani and Savona (2014). These services tend to concentrate in more densely populated regions with strong functional specialization and also should provide stronger economic growth. To achieve this, logistic regression will be applied on the resulting

clubs using the specialisation in Business services as one of explanatory variables and membership in the clubs as dependent variable. The approach taken here will be similar to Bartkowska and Riedl's, however testing directly Business services contribution and using different dataset in terms of both time period and the regions included. In this thesis the Meliciani and Savona's empirical findings about the BS concentration will be compared to the convergence clustering results.

Expected Contribution This work aims to bring light into the convergence process in the European Union. It should, at first place, be able to decide whether the convergence is present in the post-crisis situation and if yes, for which regions do we observe it and how the convergence pattern changed in comparison to previous years. Furthermore, it aims to connect sectoral specialization in different regions with relative convergence of the regions and to check robustness of convergence clubs composition by combining several methodologies for detecting them. Findings of such analysis are supposed to contribute to understanding of drivers of convergence in context of developed countries and could indicate ideal policy for achieving high convergence of the CEEC to the old EU countries. This work will therefore contribute to the literature dealing with regional convergence by analyzing convergence using the most recent dataset available and recently developed technique for detecting the clubs.

Secondly, this work will contribute to the stream of literature seeking to find factors of economic growth. Above all to the New Economic Geography (NEG) and the New Growth Theory (NGT) that focuses on endogenous factors that drive the regional development. Last but not least, it will supplement literature dealing directly with the Business services by analyzing their contribution to the economic growth empirically using recent econometric methods.

Outline

1. Introduction
2. Theoretical background – the main convergence theories introduced
3. Business specialization in the EU regions and its determinants described
4. Specification of the models used to obtain convergence clubs and degree of BS specialization
5. Discussion of the data used to obtain the model – most likely Eurostat will be used as a source of the data
6. Discussion – verification of the hypotheses using the estimated convergence clubs and estimated degree of business services specialization

7. Conclusion

Core bibliography

Arbia, Giuseppe, Julie Le Gallo, and Gianfranco Piras. "Does Evidence on Regional Economic Convergence Depend on the Estimation Strategy? Outcomes from Analysis of a Set of NUTS2 EU Regions." *Spatial Economic Analysis* 3.2 (2008): 209-24.

Phillips, Peter C. B., Donggyu Sul. "Some Empirics on Economic Growth under Heterogeneous Technology." *Journal of Macroeconomics* 29.3 (2007): 455-69.

Phillips, Peter C. B., Donggyu Sul. "Economic Transition and Growth." *Journal of Applied Econometrics* 24 (2009): 1153-1185.

Borsi, Mihály Tamás, and Norbert Metiu. "The Evolution of Economic Convergence in the European Union." *Discussion Paper - Deutsche Bundesbank* 28 (2013).

Bartkowska, Monika, and Aleksandra Riedl. "Regional Convergence Clubs in Europe: Identification and Conditioning Factors." *Economic Modelling* 29.1 (2012): 22-31.

Eckey Hans-Friedrich, Türk Martin. "Convergence of EU-Regions. A literature report." *Investigaciones regionales*, Vol.10 (2007).

Meliciani, V., and M. Savona. "The Determinants of Regional Specialisation in Business Services: Agglomeration Economies, Vertical Linkages and Innovation." *Journal of Economic Geography* 15.2 (2014).

Author

Supervisor

Chapter 1

Introduction

The aim of this thesis is to provide novel evidence on convergence of incomes across the regions of the European Union, with a focus on the development of convergence after the Great Recession. The topic of economic convergence is interesting for several reasons. First of all, convergence in standard of living among the regions and countries in the European Union can be considered desirable, as it can promote socio-economic homogeneity in the union and thus pave the way for political consensus among the member states with different historical experiences. This itself makes investigation of convergence an interesting topic for empirical research. What makes convergence an attractive object of further investigation, from the viewpoint of empirical economics, is how multifaceted and ambiguous the whole topic seems to be. For example, we can notice conflicting insights found by the empirical research in different periods of time and at the regional instead of the country level.

More concretely, despite the previous rapid development in the former, the recent development does not speak unequivocally for convergence of the Central and Eastern European countries (CEEC) to the old EU countries. Even more complicated relationship can be then 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 (Basile *et al.* 2001; Eckey & Türk 2007; Zarotiadis & Gkagka 2013).

Furthermore, the previous empirical research in this field has suggested several distinguishing features of convergence in the EU. It is often pointed out that economic development in some large urban areas, and above all capital cities, is found significantly more dynamic than development in the rest of the

countries, leading to high economic inequality within the EU states (Geppert & Stephan 2008). This seems to be especially valid for the new member states with the richest urban regions from the CEE countries, mainly areas where capital cities are located, being already among the richest regions by GDP per capita in the entire EU, while many other regions lag behind and overall discrepancy between old and new member states diminishes rather slowly. The result is that internal inequality in the CEE countries is considered especially high (Geppert & Stephan 2008; Smetkowski & Wójcik 2012; Smetkowski 2013).

Moreover, at the country level, we can notice a general shift in convergence dynamics after the 2008 financial crisis. Whereas in the pre-crisis period the CEE countries had enjoyed rapid convergence fueled mainly by strong total factor productivity growth and capital accumulation, after the crisis a substantial slowdown in convergence to the old member states took place. With the CEE countries now facing considerable external and internal hurdles to their future growth, major structural reforms are being suggested in order to re-activate the convergence process (IMF 2016). Therefore, it seems important to investigate convergence development among regions in the EU after the crisis and compare it to the above mentioned development at country level, and to compare the current convergence pattern to conclusions of the previous research on this theme. This seems important even more given the fact that most of the literature dealing with this issue uses pre-crisis data (Eckey & Türck 2007; Bartkowska & Riedl 2012).

The aim of this thesis is to investigate the dynamics of income convergence in the EU and discover presumably complex convergence pattern at its regional level using advantages of recently developed method by Phillips & Sul (2007). The Phillips & Sul's test is relatively widely used to analyze convergence at various levels and its design seems to be especially suited for investigation of club convergence that can be expected among EU regions. However, as discussed below, the test has been rarely used with recent data, whereas in this thesis, we use the post-2008 data on GDP per capita when the convergence dynamics of the CEE countries as a whole differs markedly from the previous period.

Apart from the convergence estimation itself, there is always an interest in explaining this phenomenon by some underlying factors. This thesis will examine the impact of wide range of variables, suggested in literature dealing with the topic of economic growth, on the current state of convergence in the EU. It will, however, mainly concentrate on Business services, relatively broadly de-

defined category of services including computer related activities, R&D services, legal, financial and management consultancy, advertising and marketing services among others. There is a growing literature dealing with these services and they were chosen as the main factor of interest as they represent a dynamically developing sector which can be expected to have an increasing impact on the real economy. Moreover, the literature often concludes that Business services have strong positive relationship with regional innovative potential (Kox & Rubalcaba 2007; Corrocher & Cusmano 2014; Shearmur & Doloreux 2009). We therefore assume that specialization in Business services can have significant impact on regional convergence performance.

The territorial distribution of Business services (BS), as described by Melicani & Savona (2014), shows on the one hand tendency of the services to concentrate in large urban areas, but on the other counterbalancing centrifugal forces leading away from large cities. At the same time, it is often found in the literature on economic growth that capital cities and urban areas grow faster and are more likely to enter higher Convergence Clubs. The partly overlapping nature of BS specialization and concentration of economic growth asks for deeper investigation and verification of existence of statistically significant connection between specialization in Business services and economic growth. This thesis therefore aims at investigating not only the current state of convergence pattern present in the EU but also attempts to explore the relationship between convergence and the factors that contribute to it while giving particular attention to sectoral specialization in Business services. Hence, we explore to what extent the pattern of convergence can be, among other factors, explained by the relative strength of the sector of business services.

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 sectoral specialization and other socio-economic characteristics of the regional economies in the EU and overall economic performance experienced by a certain region in comparison to the others. Particular emphasis will be given to Business services but other variables suggested by previous research will be also investigated.

The methodical approach adopted in this thesis follows the methodology used in Bartkowska & Riedl (2012), i.e., first, Phillips & Sul's log t test to find convergence clubs will be applied and then logistic regression with Business services specialization as one of the explanatory variables will be used to find

out whether higher specialization in BS increases probability of belonging to a higher convergence club. The most recent data available will be used for the analysis of convergence in the post-crisis situation. Findings of such analysis are supposed to contribute to understanding of drivers of convergence and could indicate an ideal policy for achieving fast convergence of the CEE countries to the old EU countries.

The structure of the thesis is as follows: first chapter (Literature review) describes concepts of β and σ -convergence together with the other commonly used methods for estimating convergence and convergence clubs. Then it presents the current state of the empirical research devoted to convergence and regional convergence in particular. Furthermore, it discusses some common findings about regional convergence in the CEE countries. Last but not least, Business services are defined and discussed in this chapter, mainly their composition, territorial distribution and expected contribution to economic growth as found in the previous literature. Second part (Methodology) describes in details the baseline method chosen for this thesis - the log t test developed by Phillips & Sul (2007). It also discusses the issue of spatial correlation and filtering used in this work and presents in detail the methodological framework used in this thesis for further analysis by logistic regression. Another part (Data) discusses the data we work with in the thesis, their sources, variables used and the strategy adopted for missing observations. Remaining chapters of the thesis are devoted to discussion of the estimation results. First, we discuss the composition of the convergence clubs. Later, following the logistic regression results, the role of Business services as well as other variables in regional convergence pattern is discussed. Chapter six then concludes.

Chapter 2

Literature review

2.1 The convergence hypothesis and modelling frameworks for its empirical measurement

Convergence, as well as more specifically regional convergence, is certainly one of the most discussed and empirically investigated problems of the current growth economics. In his often cited work, Galor (1996) describes three types of convergence. First of them, the absolute convergence, is defined as a long-run convergence of per capita incomes that takes place independently of the initial conditions (i.e. regardless the initial level of output). Conditional convergence, on the other hand, states that only such regions that have structural characteristics (such as technological progress, preferences, population growth, etc.) similar enough converge regardless of their initial conditions. Third option is the club convergence hypothesis according to which per capita income of two regions converge if they have same structural characteristics and at the same time their initial conditions are in a basin of attraction of the same steady state equilibrium (Galor 1996). In other words, the difference between the club and conditional convergence hypothesis can be stated as a question whether there is only one globally stable or multiple locally stable (as is the case of club convergence) steady state equilibria. One of the differences between the two concepts worth mentioning is that in case of the club convergence hypothesis transitory shocks can affect the economic performance forever which is not the case for the conditional convergence hypothesis.

As Galor (1996) notes, the absolute convergence hypothesis has been successfully empirically refuted. This can be mainly explained by the fact that the neoclassical growth model framework requires the structural characteris-

tic to converge as the country's (or region's) long run equilibrium depends on them. He states that it is somewhat paradoxical that the empirical refutation of the absolute convergence hypothesis led to rejection of the neoclassical growth framework among some researchers and development of the endogenous growth literature, as the neoclassical growth model incorporates the conditional convergence hypothesis as well as the club convergence hypothesis.

One of the empirical frameworks that builds on foundations of the neoclassical growth model is done by Phillips & Sul (2007) and will be described later in more detail. Phillips & Sul (2009) also provide somewhat different and more technical definition of convergence. They define a relative convergence as $\lim_{t \rightarrow \infty} \frac{\log y_{it}}{\log y_{jt}} = 1$ for all i and j . However, as noted by Phillips & Sul (2009), this type of convergence implies growth convergence in discrete time series and not necessarily level convergence, which is formally defined as $\lim_{t \rightarrow \infty} (\log y_{it} - \log y_{jt}) = 0$. The null hypothesis of the log t test, elaborated in detail below, primarily examines relative convergence, yet it nests level convergence. Which one of them can we actually see is then determined by value of the test's parameter γ .

Eckey & Türk (2007) bring comprehensive report of methods used for investigation of economic convergence as well as their empirical results. Probably the most common form of the convergence test is relatively simple β -convergence which departs from the transitional income dynamics equation 3.1 and in its basic version has the following form:

$$\log y_{it} = a_0 + (1 - a_1) \log y_{it-1} + u_{it} \quad (2.1)$$

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 and thus forms the conditional convergence model (Sala-i Martin 1996). It is worth noting that the β -convergence is designed in such a way that it tests the sign of the convergence parameter β in equation 3.1 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 & Sul (2007) used in this work.

Other relatively 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

$$\sigma_t^2 = (\sigma^2)^* + (1 - a_1)^2 [\sigma_{t-1}^2 - (\sigma^2)^*] \quad (2.2)$$

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, in recent years, however, many other methods were utilized for finding convergence at both state and regional level (Eckey & Türck 2007). Following part of this thesis presents a brief description of some of these methods together with the results obtained with the use of these methods.

2.2 Empirical analyses of convergence among EU regions

Eckey & Türck (2007) analyze multiple empirical works that use some of the various methods used for analyzing convergence at regional level. Among those are also σ -convergence and several specification of β -convergence. In case of β -convergence Eckey & Türck (2007) note, besides its absolute version, results of conditional β -convergence, different strategies for dealing with the spatial dependence between regions in the β -convergence framework, separate convergence analysis for different sectors of the economy or panel data version of β -convergence model.

Eckey & 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 ranging from 1980s to early 2000s. This finding is, however, not universal with some surveyed papers finding high rate of convergence, whereas others failing to find convergence at all. This variety of, sometimes conflicting, conclusions can be also caused by differences in methods as well as in number of regions included. Remarkable for this thesis are then especially limitations experienced by many authors in

case of convergence analysis of the CEE countries, caused mainly by low data reliability before and in the first half of the 1990s, which practically excludes them from any historical analysis (Eckey & Türck 2007).

Some authors also suggest that convergence among the European regions was not stable in the post-war period. Eckey & 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 then concluded that convergence in Europe appears to be U-shaped, as mentioned by Basile *et al.* (2001) and Geppert & Stephan (2008). More concretely, it seems to reach its lowest point around the beginning of the 1980s, accelerating again later on. Geppert & Stephan (2008) also conclude that convergence is happening mainly at the national level whereas regionally we do see strengthening of metropolitan areas - they however do not find the causes for this. From the newer contributions, Smetkowski & Wójcik (2012) find, using the beta convergence, only a weak tendency for regional convergence in the CEE countries, while in case of some countries (e.g. Poland) there was absolute divergence found even when the capital region was excluded. In most of the other cases, the convergence was stronger when 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 & 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 & Ord (1992) is then used as this variable most often. One of the distinguishing features of LISA is that it is aimed at finding clusters of neighbouring regions. In result, when applied to the old EU regions, this method quite often leads to large clusters along the south-north axis (as found for example in Baumont *et al.* (2003)). After accession of the CEE countries this division moves between regions of the new member states on one side and regions of the old members on the other (Eckey & Türck 2007). However, this result is not uniform, as some authors find clusters of highly developed regions across these divisions using this method (Eckey & Türck 2007).

Another method used for the detection of convergence clubs is the analysis of probability density functions of GDP per capita or labour productivity by the

Kernel density estimation (usually, the Gaussian Kernel is used). Existence of convergence clubs is then supported if there is multimodality in the probability distribution. When applied to the Western European regions for various periods of time, results of this method ranged from a single club (single peak of the income distribution) to two convergence clubs (Eckey & Türk 2007).

Furthermore, Markov chains are often used to analyze convergence clusters. This method is based on the examination of probability that regions will move from one of predetermined income clubs to another. Despite the theoretical criticism given to this method, mainly focused on its restrictive assumptions and lack of theoretical support, there are some interesting findings, especially repeatedly found evidence of poverty trap between rich and poor regions of the old EU (Eckey & Türk 2007).

The convergence literature and the range of methods listed in this short paragraph is far from complete. While we concentrated rather on the methods used most often, there are many other techniques being applied and developed. Among the newest alternative methods adopted for analyzing the convergence clubs in Europe is the approach based on Bayesian methods as suggested by Fischer & LeSage (2015) who use the Bayesian probabilities for assigning the EU regions into two predefined clubs utilizing a general stochastic space-time dynamic panel relationship between growth rates and initial levels of income. We can also see the clustering methods being employed to find convergence clubs in the regional setting. An example of this approach can be Maasoumi & Wang (2008) who use hierarchical clustering with the Granger entropy measure to analyze convergence of Chinese regions.

One of the recently developed approaches to convergence analysis is then the log t test by Phillips & Sul (2007). This method has deep theoretical foundation and nests absolute and club convergence as well as the distinction between level and growth rate convergence. As we finally decided to adopt this method for our own estimation, we devote more space to its detailed description as well as the discussion of its advantages in the Methodology part of this work.

2.3 Stylised facts about convergence in CEE regions

As already mentioned, the results of convergence analyses are various, moreover the works mentioned above often use older data (usually ending in the early

2000s as noted by Eckey & Türck (2007)). This is also true for works using the Phillips & Sul's convergence test such as Bartkowska & Riedl (2012) or Pinho *et al.* (2010). Both these works use datasets ending before 2008. Despite this shortcomings, we can provide some commonly found conclusions from recent works in the field, with special attention given to CEE countries.

Bartkowska & 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 nation to cluster together. Bartkowska & Riedl also found that regions that belong to the same club often cluster together (i.e. they are spatially correlated).

Another 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 & Wójcik (2012); Szendi (2013); Chapman *et al.* (2012); Monastiriotis (2011). Smetkowski & 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. He also notes that growth in productivity disparities was significantly lower than growth in GDP per capita disparities. Moreover, Monastiriotis (2011) finds evidence that the development in the CEE countries is formed by circular self-enforcing growth that finally strengthens the polarization effect in this region. Persistent regional disparities in the CEE countries, even in composite measures of well-being (Human Development Index), are confirmed also by Benedek & Kocziszky (2015). They found no signs of σ -convergence between 1995 and 2000.

The findings presented above show complexity of economic development in both old and new EU members and suggest that concentration on performance

at the country level can significantly blur and underestimate the real dynamics hidden behind it, at the regional level. This was the main reason why we decided to perform our analysis at the regional rather than the country level.

2.4 Business services - introduction and definition

Business services, the impact of which on the economic well-being of the EU's regions is one of the two main topics of this thesis, represent a category of services that experienced rapid growth in recent years, especially in terms of employment (Kox & Rubalcaba 2007). Business services are mostly characterized by their Business-to-business nature and by relatively high technological and knowledge content. They can also be described as activities that enhance efficiency and quality of their customers' production activities. Business services are relatively broad category involving 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 & Rubalcaba 2007; Hertog 2000). European commission (in Regulation No. 295/2008) then describes the Business services as covering a part of publishing activities, the information technology service activities, a part of information service activities and professional, scientific and technical activities as well as the employment activities.

Business services became in recent time a key part of modern economy. They 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 & Rubalcaba 2007). Their productivity growth reached somewhat worse results though, according to Kox & Rubalcaba (2007). On the other hand, Pylak & Majerek (2014) note that level of productivity is higher in business services sector than for manufacturing firms. At the same time, Pylak & Majerek (2014) finds positive correlation between share of Knowledge-Intensive Service (KIS) (more general category of services including knowledge intensive BS) and GDP per capita level.

According to Kox & Rubalcaba (2007), we also cannot consider BS nowadays as simple outsourcing of previously in-house activities. They can rather be viewed, due to the extend of product innovation and specialization they facilitate and that often change the nature of the original service itself, as a new development phase in the division of labour. Relating to this, BS have been found playing a significant role in innovation as well as knowledge diffusion -

mostly in form of spreading the "best practice" information among other companies and thus increasing their efficiency (Kox & Rubalcaba (2007); Corrocher & Cusmano (2014)). Desmarchelier *et al.* (2013) then consider Knowledge Intensive Business Services (KIBS) that represent a subset of Business Services, arguably comprising the more innovative parts of BS¹, as a category of services overcoming the stylized distinction between productive activities on one side and stagnant services on the other. Desmarchelier *et al.* (2013) stresses above all active role of KIBS in innovation process as such and in other sectors' innovations in particular. Last but not least, BS companies provide specialist knowledge to small and medium-sized firms, which would these otherwise be unable to access (Kox & Rubalcaba 2007).

As far as direct contribution of BS to the economic growth is concerned, the impact of BS is less clear. Even though Desmarchelier *et al.* (2013) show that KIBS are a factor of economic growth in a theoretical model framework, BS have also relatively weak productivity performance which is caused by the fact that employment in BS is growing faster than its value added (Kox & Rubalcaba 2007). Empirical results on the impact of BS on aggregated productivity and growth, as summarized by Kox & Rubalcaba (2007), show however rather ambiguous picture - apart from positive results for IT and computer services there are no robust results confirming positive impact of BS, when taken as a whole. Pylak & Majerek (2014) show, using structural equation modelling (SEM), that KIS have different impact in highly and less developed regions, even though in neither they influence growth directly and their impact is indirect through innovations and productivity growth. Pylak & 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 that does not have links with the KIS strong enough yet. However, for example Corrocher & 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 & Zenker (2001).

According to Shearmur & Doloreux (2009) the KIBS serve mainly as innovation transmitters - they firstly initiate and develop innovation activities in the companies of their clients and secondly support and facilitate internal

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

innovation processes of their clients. They thus function as overall transmitters of innovations and knowledge across industries through cooperation with their clients. Apart from this role, KIBS can be thought of as active creators of innovations - even though, as noted by Shearmur & Doloreux (2009), the character of their innovation process can differ from that in manufacturing sector, which also complicates measurement of their contribution to innovations.

Specifically for the CEE countries, Baláž (2004) notes that the economic transition was likely a major factor for growth of KIBS in these countries (whereas growth of communication technologies was more important in richer countries). He also notes that convergence of the CEE countries to the OECD standard should be accompanied by further rise of this sector.

Moreover, there is an evidence that the business services show clear affinity towards large metropolitan areas (and areas with high population density in general), as shown by Meliciani & Savona (2014), when analyzing determinants of location of the BS. The same authors, however, show that this tendency in localization of business services is somehow counterbalanced by centrifugal forces away from large metropolitan and urban areas. These forces are then led by presence of such companies from the manufacturing sector that are intensive users of BS. Similarly, Guastella & Timpano (2016) find that agglomeration is not sufficient condition for economic growth with some regions experiencing expanding agglomerations and at the same time relatively slow growth. Last but not least, specialization in BS is found to be positively correlated with strong ICT as well as R&D sectors (Meliciani & Savona 2014). Guerrieri & Meliciani (2005) also find that presence of knowledge-intensive manufacturing industries has strong effect on specialization in business services and thus further confirms the link between productive industries and BS. Meliciani & Savona (2014) then show in their analysis that capital cities exert negative influence on degree of BS specialization in neighbouring regions (this can be compared to Bartkowska & Riedl's conclusion for regional clusters in Western Europe), whereas highly populated regions give their neighbours positive spillovers on specialisation in BS.

2.5 Business services as a factor of economic development

In this thesis, we shall, apart from looking for the convergence clubs, attempt to find factors that influence the economic growth and bring specific regions to certain convergence clubs. Mora (2008) mentions several factors considered in convergence literature as possible determinants for emergence of different convergence clubs. He 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 neo-classical growth model. Especially, differences in initial human capital distribution and in initial income distribution can bring otherwise structurally same economies to different equilibria. Furthermore, Galor (1996) shows, in several theoretical model settings, that endogenous fertility decisions can lead to club convergence. Interestingly, Galor (1996) mentions that perfect international capital mobility can undermine the club convergence hypothesis - but not in case when human as well as physical capital is considered. Also technological progress can in some settings lead to multiple steady state equilibria (Galor 1996). Finally, different parameters, such as saving rate, of the considered production function can lead to convergence clubs emergence (Mora 2008).

Guastella & 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 & Timpano 2016).

In a sense, the club convergence phenomena, as presented in Mora (2008), can be entirely related to local population's characteristics. Practically, this is often expressed in econometric analysis by including multiple control variables to the regression function such as saving rates, population growth variable, human capital endowment (ratio between high and low levels of education), proxy for R&D (patent application per million inhabitant). Authors also usu-

ally include sectoral specialization indexes and other variables (Mora 2008). Bartkowska & 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 & 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, described in more detail in the Data section, was strongly influenced by the limited availability of many measures on the regional level. Despite this, we managed to include a set of variables that, in our belief, to a large extent, match those used in the works just mentioned - general population characteristics, measures of research/innovation activity, sectoral specialisation, dummies meant to express institutional quality, the regions' relative geographic location and variables measuring economic attractiveness. We also tried, where it was possible, to use the variables in form of initial factors, i.e. measured in 2008.

Chapter 3

Methodology

3.1 The log t test by Phillips and Sul - empirical properties

The method that will be used in this thesis for analyzing convergence is the so-called log t test developed by Phillips & Sul (2007; 2009). This method for testing for convergence builds on the foundation of the neoclassical growth model and β -convergence in particular. According to Phillips & 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 3.1. Phillips & Sul (2007) show that this assumption leads to inconsistency of classical β -convergence stemming from omitted variable bias as well as endogeneity. In result, the estimates of the β -convergence may actually have reverse sign and thus give us completely wrong insights about the character of convergence and impact of investigated growth factors on economic performance (Phillips & Sul 2009). The log t test by Phillips & 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 Phillips & Sul (2007)'s test is relatively widely used at national level, as shown by Borsi & Metiu (2015); Fritsche & 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 & 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.

The Phillips & Sul's test was praised for its flexibility and ability to find different types of convergence paths for each cross-sectional unit. It can also distinguish between absolute and relative convergence (Borsi & Metiu 2015). This test is furthermore able to detect convergence even in situations where traditional convergence tests fail thanks to its treatment of convergence as an asymptotic property (Bartkowska & Riedl (2012); Borsi & Metiu (2015)). Moreover, the log t test was, in the above mentioned works, shown to give convergence clubs, members of which are not geographical neighbours even though such result is not at all excluded. This seems to be in contrast with the LISA analysis, described previously, which almost necessarily lead to large continuous blocks of regions. We found this property of the log t test appealing and concluded that besides mending the flaws of traditional β -convergence, it also gives us convergence clubs that can be considered much more plausible than it is the case for some of the leading methods for analyzing club convergence. All of these were reasons for choice of the log t test for our analysis.

3.2 The log t test - modelling framework

The Phillips & Sul's test represents, according to the authors themselves, an econometric perspective of theories elaborated by Lucas *et al.* (2002) who optimistically claim that large economic inequality in the world economy, typical for the period after the Industrial revolution, is coming to an end with the beginning of the 21th century, thanks to, above all, increased human capital accumulation that encourages sustained economic growth and is largely achieved via demographic transition in form of declining fertility.

Phillips & Sul's motivation is to build a test that allows for empirical study of economic transition. The resulting framework is supposed to embody such transition behaviour over time that can incorporate temporarily economic divergence as well as anticipated long-term economic convergence (Phillips & Sul 2009). As welcomed side effect, the test allows us to study the convergence patterns in high detail and precision.

The model departs from a variation of the neoclassical growth model where equal technological progress over time is assumed for all countries regardless of their initial conditions. Under such assumption it is difficult to explain the cross-section income heterogeneity that is being perceived in the real world (Phillips & Sul 2009). Following previous theoretical works, Phillips & Sul

(2009) therefore depart from a model for log per capita income that combines idiosyncratic and individual elements.

Concretely, the desired transitional cross-sectional divergence is possible thanks to the fact that parameters β_{it} and x_{it} from the following expression mapping the transition dynamics of real per capita income:

$$y_{it} = y_i^* + a_{i0} + (y_{i0} - y_i^*)e^{-\beta_{it}t} + x_{it}t \quad (3.1)$$

are allowed to vary across cross-sections. Variables y_{i0} and y_i^* in 3.1 are initial and steady state levels of log per capita income, x_{it} is component expressing technology accumulation over time and a_{i0} captures initial technology accumulation. Parameter 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 also claim that β_{it} can be reasonably considered as an increasing function of the technological progress parameter x_{it} . This fact 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 in this framework, which gives rise to differences in speed of convergence given by parameter β_{it} . As Phillips & Sul (2009) note, parameter β , in its homogeneous version, is what traditional β -convergence test is trying to estimate and assesses the convergence depending on whether this parameter is positive or negative. The authors, on the other hand, build their scheme based on heterogeneous speed of convergence.

Given the definition of technology accumulation from the previous paragraph, we assume $x_{it}t$ to contain both idiosyncratic and shared element across economies and we express the equation 3.1 as

$$y_{it} = \left(\frac{y_i^* + a_{i0} + (y_{i0} - y_i^*)e^{-\beta_{it}t} + x_{it}t}{\mu_t} \right) \mu_t = \delta_{it} \mu_t \quad (3.2)$$

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 . In other words, it is economy i 's share of the common trend μ_t and is determined by characteristics of individual re-

gion. As we will see later, parameter δ_{it} proves to be crucial for determining individual region's level of convergence.

Note that by defining $a_{it} = y_i^* + a_{i0} + (y_{i0} - y_i^*)e^{-\beta_{it}t}$, we can see that as $t \rightarrow \infty$ the time dependent element of a_{it} decays to zero and y_{it} in 3.1 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.

As can be seen above, the loadings δ_{it} , apart from already discussed parameters β_{it} and x_{it} , depend also on the initial level of technology and the steady state level of per capita income that are both contained in a_{it} . In result, the necessary condition for growth rate convergence: $x_{it} \rightarrow x$ as $t \rightarrow \infty$ for all i , is not sufficient for level convergence of y_{it} .

Moreover, if we represent the common steady state growth element μ_t by either unit root stochastic trend or simple linear deterministic trend $\mu_t = t$, as can be done in a neoclassical growth framework, we can easily see that

$$\delta_{it} = x_{it} + \frac{a_{it}}{t} \quad (3.3)$$

and thus $\delta_{it} \rightarrow x_i$ 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 $\delta_{it} \rightarrow \delta_i$ with $t \rightarrow \infty$ 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:

$$H_0 : \delta_i = \delta \quad \& \quad \alpha \geq 0 \quad (3.4)$$

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

$$H_A : \{\delta_i = \delta \quad \text{for all } i \text{ with } \alpha < 0\} \quad \text{or} \\ \{\delta_i \neq \delta \quad \text{for some } i \text{ with } \alpha \geq 0, \quad \text{or } \alpha < 0\} \quad (3.5)$$

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

$$h_{it} = \frac{y_{it}}{N^{-1} \sum_{i=1}^N y_{it}} = \frac{\delta_{it}}{N^{-1} \sum_{i=1}^N \delta_{it}} \quad (3.6)$$

This formula traces trajectory of each cross-section i relative to the club's average and is thus called by Phillips & 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} :

$$H_t = N^{-1} \sum_{i=1}^N (h_{it} - 1)^2 \quad (3.7)$$

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 convergence to zero with time going to infinity. If it remains positive, we conclude that convergence did not happen. Interestingly, as H_t typically converges to a positive constant in case of club convergence, the test is suitable for convergence clubs recognition, as we will see below (Phillips & Sul 2009).

The econometric 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 & 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: $H_t = \frac{A}{\log(t)^{2t^{2\alpha}}}$ as $t \rightarrow \infty$, which then leads to the final formulation of log t regression:

$$\log\left(\frac{H_1}{H_t}\right) - 2\log(\log(t)) = a + b\log(t) + u_t \quad (3.8)$$

This test is called log t convergence test by the authors and has one sided t -test test of hypothesis of convergence against no or partial convergence in

its core, as shown above. The 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 framework above can be used to test for absolute convergence across all countries as well as for detecting the club convergence if the absolute convergence is not confirmed. 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 as defined above (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.

3.3 The log t test - formation of convergence clubs

If the hypothesis of overall convergence is rejected, there comes a procedure for checking whether there is not convergence in subgroups of the investigated sample. Phillips & Sul (2007) developed a special testing algorithm for this purpose that works with repeated log t regressions. This algorithm is formed of several steps. First, core group of countries is localized by maximizing the log t test statistic over different sizes of subgroups of the k highest observations. Then, if the test rejects the null of convergence for the two highest observations of the subgroup, the highest observation in the group is dropped. The entire procedure is then repeated until it is not possible to drop the highest observation according to the above mentioned principle. Subsequently, remaining observations are tested for membership in this core group and alternatively the process is repeated from the beginning for the remaining observations. If no converging individuals are found, we conclude that there is no convergence in the sample. In detail, the clustering procedure suggested by Phillips & Sul (2007), below described at country level, follows these steps:

- Order the countries according to the amount of last period income.
- 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 : $k^* = \operatorname{argmax}_k \{t_k\}$ 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 then compared to criterion level c^* . In our case we choose $c^* = 0$. If the associated t -statistic is greater than c^* , we add the country 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.

We could notice choice of the criterion $c^* = 0$, in the second step of the procedure. This value 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 & Sul (2009) note that c^* can, in general, 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 possess 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 & Sul (2009) and further elaborated in detail by Bartkowska & 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 the 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.

3.4 Spatial correlation - issues and solutions

According to Dall'Erba & Le Gallo (2008) were most of the empirical tests for regional convergence based on the same assumption as the tests for the international income convergence. This means that regions were considered to be independent units and any potential links among their economic activities we ignored. Magrini (2004) adds that the neoclassical model, on which β -convergence is based, works with an assumption of a closed economy. Although this assumption may hold in case of entire countries, it is highly unreliable for regions, which are often, above all, administrative units without any barriers to trade comparable to those at the country level (Magrini 2004). This leads to so-called spatial autocorrelation.

There are two versions of the spatial autocorrelation, which can be generally defined as spatial dependence between geographically organized units - so called substantive, caused by existence of actual interregional linkages and nuisance, caused by a measurement error (Magrini 2004). Getis & Ord (1992) compare spacial autocorrelation to temporal autocorrelation, which is common and well studied in time series. Same as for time series, the main problem associated with spacial autocorrelation is correlation between model's errors leading to model misspecification.

Anselin & Rey (1991) note that spacial autocorrelation of the kind that we can expect in our case (i.e. autocorrelation stemming from mismatch between real economic activity and borders of administrative units for which the data are collected) should not bias the estimates of OLS regression but it does so in case of model's variance, which leads to biased t -tests and measures of fit. Given the previous research such as Bartkowska & Riedl (2012), we can expect autocorrelation in logarithm of GDP per capita among the European regions as well. We indeed confirm existence of the spatial autocorrelation in our data by Moran's I test statistic¹ being equal to 0.53 and the associated test clearly rejecting the null hypothesis of no spatial correlation. In sum, the spatial autocorrelation among the EU regions in log GPD per capita could harm the

¹Moran's I is a statistic used for measuring the spatial autocorrelation and an alternative to Getis & Ord's statistic that is used in this work. It is constructed as a two-dimensional version of univariate correlation test for time series (Anselin 2001).

results of Phillips & Sul's log t test, as the key parameter b is based on an OLS regression, in which the dependent variable is function of the log GDP per capita.

The current literature on regional convergence, such as Bartkowska & Riedl (2012), use various methods to deal with this type of autocorrelation. One of these methods, developed by Getis & Griffith (2002), will also be used in this thesis. Getis & Griffith (2002)'s filtering approach is 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. In practice, the filtering follows formula:

$$x_i^* = \frac{x_i \frac{W_i}{n-1}}{G_i(d)} \quad (3.9)$$

Where n is then number of regions within the distance d from region i . W_i is defined as sum of all regions within distance d . Technically, W_i is sum of all geographic connections w_{ij} , each of which takes value 1 if unit j is within distance d from unit i and 0 otherwise. $G_i(d)$ is value of Getis and Ord's statistics measuring degree of spacial association for given region within distance d . This statistics is defined by Getis & Ord (1992) as

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_j x_j} \quad (3.10)$$

Crucial part of the whole process is choice of optimal distance d . Following Getis & Griffith (2002), we evaluated the G_i statistics for increasing distances d until the we could find no further autocorrelation, or in other words, until the value of G_i statistics started to decrease. Despite the fact that we expect the character of spatial interaction in Europe to be relatively stable over the given period of time, we calculated optimal d separately for each year in our dataset. The filtered variable x_i^* , log GDP per capita in our case, is then used in the test designed by Phillips & Sul (2007) and described above.

3.5 Impact of the Business services - methodology

The key goal of this thesis is an investigation of, presumably positive, impact of region's specialization in Business services on its economic performance. In the Literature review we summed up existing research on Business services and noted several arguments why this impact can be significant and positive. In

this section we will continue by specifying framework we used in order to test whether our assumptions were correct.

As already described, referring to work of Meliciani & Savona (2014), the concentration of the Business services is split between large urban areas and regions with high presence of specific industries demanding them. Additionally, despite being shown by Guastella & Timpano (2016) as not unequivocal, it was often concluded in the growth literature that large urban areas grow faster and are likely to enter higher convergence clubs. These irregularities point at complexity of the regional economic dynamics that represents the main inspiration for a thorough analysis of the factors driving the economic convergence and particularly for verification of existence of statistically significant connection between specialization in BS and economic growth.

To achieve this, first, the test by Phillips & Sul (2007), described earlier, will be used to obtain actual composition of the current convergence clubs. Second, approach inspired by Bartkowska & Riedl (2012) will be applied to the previously found convergence clubs. This means that logistic regression model will be used with a set of explanatory variables that are likely to influence the club membership. Special attention will be given to specialization in Business services as an explanatory variable in order to find out whether higher specialization in BS increases probability of belonging to higher convergence club.

As mentioned above, the conclusions found about the convergence pattern in the EU are not very robust, and there is evidence for change in the convergence pattern over time. Also we may observe in the research papers on economic convergence that most of them use older, namely pre-crisis, data with the most relevant paper Bartkowska & Riedl (2012) using data between year 1990 and 2005. Therefore, among the goals of this thesis will certainly be the utilization of the most recent data with an aim to investigate the state of convergence and compare composition of the clusters to those found earlier. Given the variety of conclusions about convergence clubs, it would be certainly beneficial, similarly to Arbia *et al.* (2008) in case of absolute convergence, to use one or more alternative clustering mechanisms (such as LISA or Cluster analysis) on the same dataset and compare the empirical results of these analyses in order to evaluate robustness of the conclusions reached using the Phillips & Sul's framework. Yet such task is out of scope of this thesis.

Chapter 4

Data and Variables

4.1 Explanatory variables

The data used in this thesis were obtained primarily from the Eurostat database¹. For the convergence club detection the log of GDP per capita was used as suggested by Phillips & Sul (2009). We worked with data from 2000 to 2015 whereas most of the explanatory variables, mentioned below, were used only since 2008. This is mainly due to time limited availability of these data for the new member states as well as changes in the Eurostat definition of Business services as discussed below. On the other hand, this setting allows us to fully concentrate our analysis on the post-crisis development.

Among the aims of this thesis is also to contribute to literature dealing with the impact of Business services on economic development - the Business services are therefore crucial explanatory variable in the model. We found two ways of expressing the level of specialization of an economy in BS. The first framework was suggested by Meliciani & Savona (2014) who define specialization in BS as a location quotient defined as employment in the Business services in certain region over total employment in that region divided by employment in the Business services for all regions over total employment for all regions. Alternatively, Guastella & Timpano (2016) use in their work somewhat simpler measure, defined as a share of workers employed in the Knowledge Intensive Services in total employed population. Due to limited data availability we stick to the second option in this work, while we used BS instead of KIS.

In the logistic regression, mentioned below, we work with a change in the specialization in Business services between years 2008 and 2014. This particular

¹Concretely, we mostly used the "Regional statistics by NUTS classification" database.

time span was determined mainly by changes in Eurostat's definition of BS as well as the data availability in case of some, especially new, regions of the EU. The same time span was also preferred for all other explanatory variables.

This work aims to 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). In practice, this 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. This is certainly not an ideal situation. However, as noted by Toivonen *et al.* (2004), definitions of BS or in practice very similar category of KIBS, are various and they differ from author to author (Meliciani & Savona 2014). Interestingly, most of them include R&D in some form, R&D are however not directly included in the definition used by Eurostat. According to Toivonen *et al.* (2004) most of the KIBS definitions contain computer and related activities, R&D services, legal, financial and management consultancy, advertising and marketing services, and technical services. More detailed definition of Business services can be found in the Literature review.

Apart from the change in Business services specialization, we used several other explanatory variables for our analysis. The choice of these variables follows New Economic Geography (NEG) and New Growth Theory (NGT), as mentioned earlier, as well as the choice of variables in some other, similar, research papers such as Mora (2008) or Bartkowska & Riedl (2012). Table 4.1 presents an overview of these variables with their names as they appear in the regression results below.

First of all, we included some general population characteristics. Concretely, among the variables are population growth, use of which was inspired by neo-classical theory and the variable was often used in similar frameworks, 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 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 &

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 also 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 two variables measuring level of specialization in other areas apart from the BS - namely agriculture and industry measured as share of population employed in these sectors over the overall employment in all areas for year 2008. Moreover, we constructed several other variables directed at measurement of either institutional quality or level of urbanization and economic concentration in the regions. Concretely, we used percentage of individuals that have an interaction with public authorities via the internet. We believe that this variable represents 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).

4.2 Treatment of missing observations

Before further econometric analysis of the factors promoting economic growth, it is necessary to find a solution to the issue of incomplete data, which is especially valid in this work for the Business services but also affects most of the other variables. It holds for all the variables used that, in the Eurostat database, the data in case of the new member states are basically available only since the year when the particular country entered the EU. This would exclude most of the Central and Eastern European countries from the analysis before 2004 and Bulgaria and Romania before 2007. Secondly, the key variable of interest, the Business services, is 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, also a change

Table 4.1: 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) |
| Infrastrurcture | 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.

in the definition of BS itself. At the same time, the data collected on Business Services after 2008 are not perfectly in accordance with the EU's definition for that period (as mentioned above). We therefore considered these two series comparable only with difficulties. For all these reasons, we collected only the data since the year 2008.

Another problem is that the data are incomplete for several regions across the explanatory variables. Given the importance of the variable for our research question, we examined Business services in the biggest detail. Concretely, for the Business services we have, from 275 NUTS 2 regions altogether, considered 184 regions with complete data. What is typical for this phenomenon is that large part of these missing regions are actually not of much interest for this work (small islands, or other remote overseas territories). However, there seems to be problematic regions, or more precisely problematic countries, even on the European mainland.

We can see a pattern where a region that once was unavailable tends to be unavailable also in at least some of the remaining periods and these regions tend to be located within several countries (Greece, Denmark, Belgium, two regions in Germany, systematic in this regard is France where data are available only since 2010). Also, there have been some changes in the NUTS 2 units during the measured period with several regions undergoing territorial changes (namely 3 German, 2 Italian, all Finnish regions apart from one and also two UK's regions) making their comparison questionable. However, as these territorial changes tend to be rather small and given the negligible size of the change of the variable of interest (BS employment) caused by these changes, we can conclude that it is possible to ignore them. Furthermore, the London area and Brandenburg were affected by a changes at NUTS 2 level, these were however simply divisions (or merges) of the previous NUTS 3 units without any boundary shift. Luckily, as far as the CEE countries are concerned the data availability is much better than for the old EU countries. Namely, for the 2008-2014 time period there are all the data available for all the CEE countries with an exemption of two regions in Poland in the year 2012 and two in Slovakia in 2009. What is also important is that in case of regions that contain large metropolitan areas the data availability is in general good across the whole sample.

As far as the Business services are concerned, we can see 184 NUTS regions with complete data for the entire period. Another 36 regions were interpolated using geometric mean of employment in BS between years 2008 and 2014 as a measure of growth rate of this variable. Another 25 regions with more than

3 missing BS values were interpolated using several interpolation methods detailed below.

Finally, 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, or combination of both - we found such data unsatisfactory for estimation of the convergence dynamics.

We used several methods to interpolate the missing values. Most of the explanatory variables were chosen with respect to their completeness, i.e. when the desired variable had many missing observations we tried to find similar and more complete one. The problem of missing values was then further alleviated by using average values for the given period. The remaining missing observations were imputed using three different methods for interpolation.

In case of Business services we, first of all, alleviated the need for interpolation by concentrating on change in level of specialization in BS between 2008 and 2014, making the intermediate values relatively unimportant. In case of regions that only have one or two missing observations, and these were at the end or the beginning of the observed period, we interpolated the missing values using average growth of BS between the first and the last year that are not missing and interpolated the missing values, i.e. years 2008, 2014 or both, by multiplying or dividing the nearest value by the average growth.

For the remaining 25, mostly West European, regions with missing values for BS, we considered using similar form of imputation while utilizing older data between years 2000 and 2007 (during which Eurostat used slightly different even though in general comparable definition of the BS) to obtain the estimate of average growth of specialization in Business services. However, as there is the already mentioned change in the BS definition and, perhaps more importantly, the average growth of the BS seems to be significantly higher before than after the 2008 financial crisis, we did not consider this method to be completely reliable and turned to alternative imputation methods to ensure robustness of our results.

Specifically, we used three other methods. All of them share a common feature, namely the fact that they consider the entire matrix of explanatory variables to estimate the missing observations that are subsequently imputed.

In particular, we used k-Nearest Neighbours algorithm that works with the proximity between the observations as measured using the complete variables. The missing values are then estimated as an weighted average of the k nearest

neighbours' values for the respective variable (we used $k = 10$). As there is relatively low number of missing observations we believe that this method can give us reliable estimates. Secondly, we used factor analysis for data imputation, a method based on the Principal Component analysis and contained in the R's package *missMDA*. We used the three most informative principal components for the imputation. 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.

Chapter 5

Results

5.1 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 & 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. As suggested by Phillips & 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, we could, first of all, see that the test clearly rejected the overall convergence among all the regions of the EU with the log t coefficient being equal to -20.57569. This result was not a surprise given the conclusions of the previous research dealing with club convergence. We can therefore see that there is no overall convergence in the EU after the Great Recession. The result also speaks against the findings of earlier studies, based mostly on β -convergence. The composition of the convergence clubs will later further suggest that we cannot speak about convergence of the old members either.

We therefore continued with identification of potential convergence clubs. However, before that we conducted a visual analysis of the convergence pattern. Following Phillips & Sul (2009), we used the parameter h_{it} , which is called relative transition coefficient and traces out individual transition path of each cross-sectional unit with respect to the average level of the dependent variable for a given group of regions. Therefore, it can be used for investigation of development of relative convergence over time. We used this parameter for graphical investigation and illustration of relative transition patterns and

stages of convergence among the EU NUTS 2 regions. Figure 5.1 below illustrates lack of overall convergence among the EU regions. It depicts relative transition paths from the overall convergence test with the analyzed regions grouped into five convergence clubs as later constructed by the log t test's club formation procedure (i.e. it shows development of average h_{it} over time for each convergence club with respect to the common limit). The curves can be thus interpreted as the respective club's performance against the overall average of log GDP per capita. We can see from the plot that there is no clear convergence pattern with almost all of the clubs staying far from the value of 1 without significant tendency to move towards it, which would be behaviour necessary for convergence. We can also note that the general relative behaviour does not seem to change around the year 2008.

Following the Phillips & Sul's clustering procedure, we initially obtained 10 convergence clubs and two diverging regions. Concretely, these were two regions with the highest level of income from entire sample - Luxembourg and Inner London – West. Neither first nor any other convergence club could be assigned the membership of these two regions according to the log t test (they are represented by the highest transition curve in figure 5.1 and as Club 0 in figure 5.4). As suggested by Bartkowska & 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 5.1 and 5.2 show results of the procedure both before and after the merging.

Table 5.1: 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

As noted above, the value of the coefficient b in the test regression 3.8, which can be 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

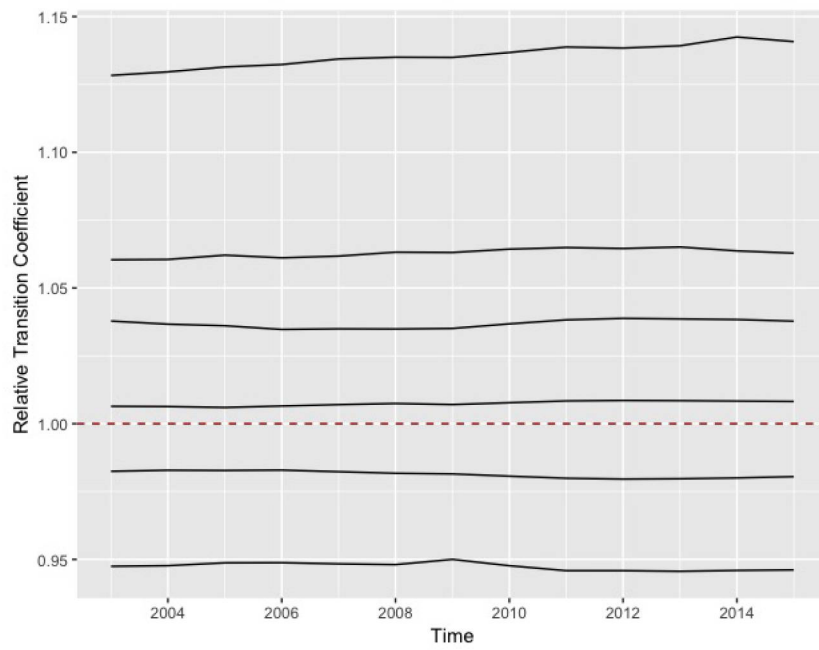


Figure 5.1: Average transition paths across merged convergence clubs

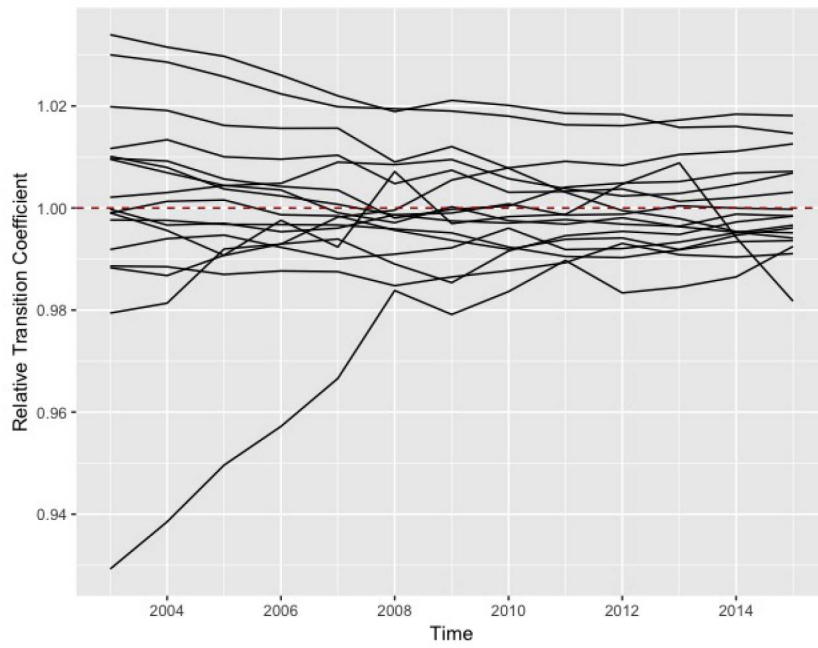


Figure 5.2: Transition paths for members of the first Club

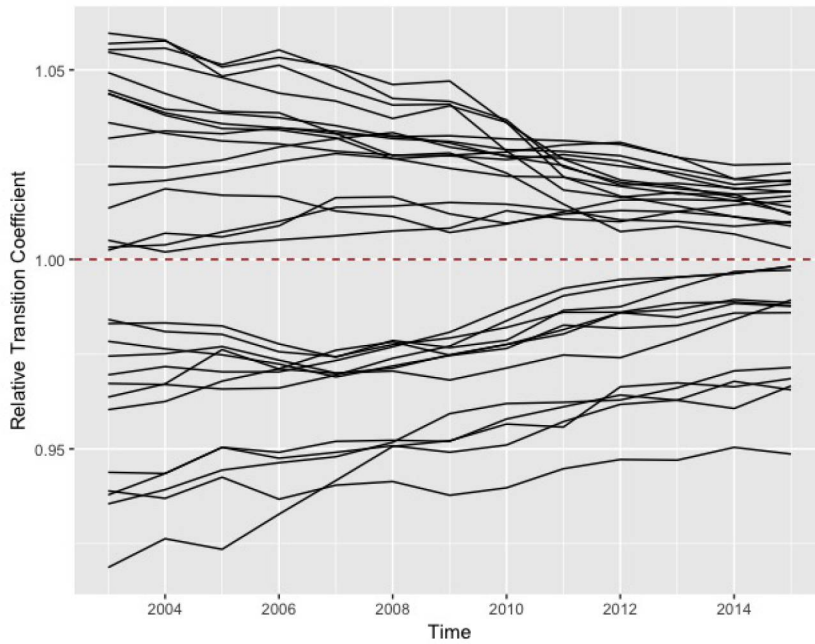


Figure 5.3: Transition paths for members of the fifth Club

the first convergence club shows strong relative (growth) convergence with significantly positive estimate of b equal to 0.26612. On the other hand, all the 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 (Phillips & Sul 2007). We cannot confirm level convergence within any of the clubs. These results are similar to those of Bartkowska & Riedl (2012) for earlier period and limited to Western Europe only.

In contrast, the clubs before merging show more robust results in terms of their convergence performance. We see higher value of t -statistics with all the b coefficients being positive, which is intuitively desired result, and several of them being even statistically significant. This suggests that choice of the conservative convergence criteria ($c^* = 0$) seems to be relatively wise. 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. Therefore, we continue our analysis with both sets of the clubs.

Besides the convergence testing by the log t test, we also plotted the convergence behaviour using the relative transition coefficient h_{it} . Figures 5.2 and 5.3 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 clubs rising from very low relative level and reaching the average level of the group, others on the contrary descending from relatively higher level of GDP per capita to the group average. We can also note changes in convergence pattern with some regions showing alternating states of convergence and divergence from the club's average.

Table 5.2: 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

Apart from their role as a formal test of convergence, we see the clubs resulting from Phillips & Sul (2009)'s clustering method as an intermediate step towards analyzing the convergence in the EU regions in more depth. Figure 5.4 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 5.4¹. These results thus confirm previous findings about convergence clubs as made for example by Smetkowski & Wójcik (2012) or Bartkowska & Riedl (2012).

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

¹Note, that there are also the outlying regions (Luxembourg and Inner London - West) included, forming the "Club 0". Also, the size of the plot does not allow us to depict the Prague region, Prague is actually at the very first club (see Appendix). Regions in the Czech Republic thus further confirm incidence of strong income inequality in CEEC.

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 & 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 & 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. At the same time, the performance of the capital cities as well as the emergence of the "Blue banana", centered around southern German regions, in the heart of the continent, also suggest an existence of broadly defined centre-periphery division as significant.

In summary, the results of the log t test and subsequent clustering procedure has proven the absence of solid convergence in the level of income among the EU regions and has further shown that there are several distinct groups of regions characterized by different economic dynamics, with centre-periphery division or potentially north-south division looking as its most plausible description. We did not confirm any structural lagging behind of the CEE regions, even though we can confirm substantial income inequality within these countries, as mentioned in the literature review.

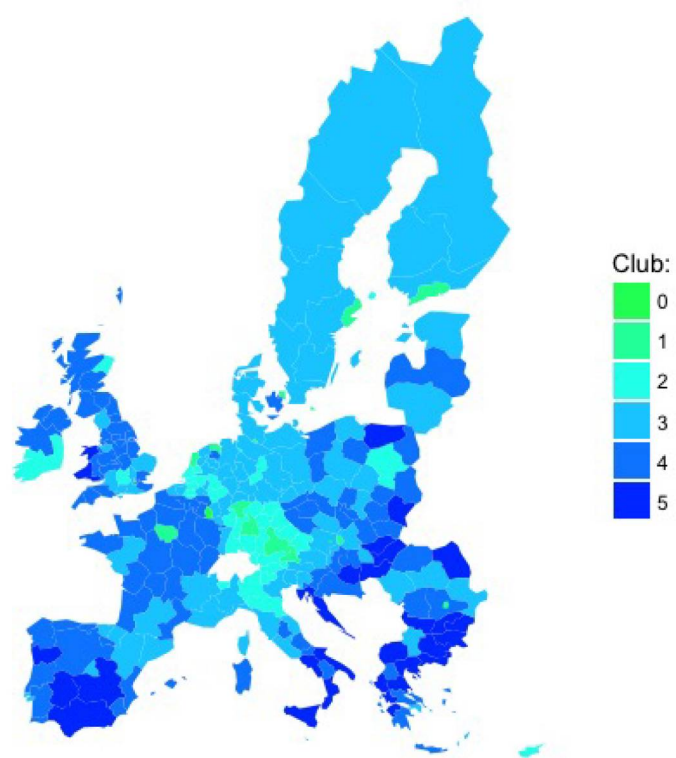


Figure 5.4: Convergence clubs and diverging regions (Club 0)

5.2 Determinants of club membership - estimation framework

After finding the convergence clubs we continued our analysis with a regression. We tried to analyze factors that influence economic performance of the European regions. In accordance with the previous research, we have chosen variables that are likely to have a significant impact on the club membership. Given the topic of this thesis, the main focus was nevertheless the change in the region's specialization in Business services. We expected positive impact of these services on regions' probability of entering higher convergence club. No specific hypothesis apart from general intuition were given for the other variables.

We departed in our estimation from the previously found convergence clubs, which we used as dependent variable. In order to confirm robustness of our results, we decided to use clubs both before and after applying the merging procedure, both these specifications led to approximately the same results (in this chapter, we show results for the clubs after merging, a version with 10 clubs can be found in the Appendix). Given the character of our dependent variable, which is discrete and can be easily and logically ordered, we considered the ordered logistic regression framework to be the best for our analysis.

As already mentioned in the discussion of the Data used, we first of all had to deal with the missing observations. Originally, we considered interpolating the missing values using the average growth of employment in BS in the period before 2008. However, due to significant change that poses the year 2008 in the dynamics of employment in Business services as well as limited data availability for some of the regions, we in the end preferred two methods of interpolation that are based on similarity between the cross-sectional units. First of all, we used factor analysis, a method based on Principal components, for imputation and we compared the results with imputation done using the k-Nearest Neighbours. Results of the regressions with the factor analysis imputation can be seen below, the results achieved using kNN are in the Appendix.

We estimated several versions of our regression. Apart from estimating the regression separately for 10 and 5 clubs and with different imputation methods, we also experimented with including the two regions that we found as diverging (Inner London - West and Luxembourg) but as expected this did not substan-

tially change the results². Furthermore, 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 & 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 (details are discussed in the Data section), resulting in totally 219 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 & 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.

Below, in Tables 5.3 and 5.4, we can see results of some of the regressions just mentioned (the rest is in the Appendix or available upon request). As usually, we show the standard regression outputs with regression coefficients and their standard errors. The regression coefficients are however at maximum informative for the extremes of the dependent variable's distribution (Carroll 2017). We therefore conducted an analysis of marginal effects (calculated at the means of the independent variables) for each of the variables. Marginal effect shows impact of one unit change in an explanatory variable on probability that

²The regression outputs are available upon request.

certain region will enter particular club. In Tables 5.3 and 5.4 we report the marginal effects, standard regression outputs are reported in the Appendix for brevity.

The regression results show that over various regression specifications the concrete values of estimated coefficients somewhat change, but in most cases we can reach similar qualitative conclusions. A more detailed comparison of the weighted maximum likelihood estimation and the not weighted results however shows that these differ in significance of some of the parameters. Despite this, we can reach some conclusions common to all of them.

Table 5.3: Marginal effects: Logistic regression with merged Clubs and FAMD^a 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

^aFactorial Analysis for Mixed Data, an imputation method based on Principal Components Analysis

5.3 Drivers of Convergence - discussion

The first thing we have noticed is a significant decrease in the magnitude of the marginal effects for the very first (and to lesser extent also the last) convergence club. In most of the cases the marginal effects are very close to zero for the highest club. 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 magnitude of the effect for the highest club. This may certainly be explained by small number of observations in the club, yet it is important to note that the marginal effect's change is disproportionately larger than difference in the relative size of the clubs. Moreover, the marginal effects for the first club turned out to be insignificant in the weighted logistic regression. This puts doubt on ability of the variables used to explain the membership in the highest club.

This conclusion is true also for our main variable of interest - 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 positive, yet significantly lower than in case of the two previous clubs. It is necessary to say though, that when we estimated our regression with weights, Business services had significant effect only on entering the second highest club even though very similar to the not weighted regression's result. This in general confirms our hypothesis about the positive effect of specialization in Business services, yet makes us doubt the robustness of these conclusions.

Apart from that, we can 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

Table 5.4: Marginal effects: Weighted Logistic Regression with merged Clubs and kNN^a 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 |
| Infrastructure | -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

^ak-Nearest Neighbours

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 even though the marginal effects' estimates gave us what we expected.

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.

In case of Business services, the results seem to confirm our expectations about their positive impact on regional economic performance. A notable aspect of our findings however seems to be the fact that they have a positive

impact on the higher clubs (second and third) but when it comes to the first club their power appears to diminish. This may be explained by rather supportive role and nature of Business services. Hertog (2000) notes that, despite the BS provider plays a key role in facilitating the innovation process, the innovation as well as conduction of the innovated activity itself is usually done by the original client³. Thus we can only hardly imagine a region creating its prosperity only via specialization in this supportive services. 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.

As already noted, the suggested lack of impact on the highest club seems to be a general phenomenon pointing to possible missing factor in our analysis that could lie behind the very top economic performance of some regions. As far as the other variables are concerned, we mostly found effects in line with our expectation/economic theory even though some of them were not significant. The most important finding then seems to be the suggested importance of applied research for regional development that can be seen from the results.

Bringing together the interpretation of both the convergence clubs alone and the just discussed regression results, 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 for a creation of the 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.

³Hertog (2000) mentions "an engineering consulting firm (that) supported an oil and gas company that wanted to drill and explore in a 'protected' area, helping them to find new operational methods to meet the strict environment protection rules by reviewing existing practice, proposing new operations, designing new methods, training the operators of the client firm, actually heading the drilling operations" as an example of such cooperation between BS provider and a client.

Chapter 6

Conclusion

In this work we aimed at a deeper investigation of convergence in the European Union after the economic crisis that started in 2008, while looking at the regional perspective rather than the country level. The regional level was chosen as it can give us a better insight about the convergence process inside the EU countries. Clear motivation for this work was then the economic performance of the Central and Eastern European members of the EU. Their tendency to strong convergence in the pre-crisis period seemed to stall during the crisis years. The choice of the regional level was also motivated by the often found evidence for strong regional economic inequalities characterized by prosperous capital cities in contrast to many regions that lag behind. Also, the ambiguity of findings of the literature about the convergence at regional level suggested a need for further research.

For this reason, we decided to apply the convergence test developed by Phillips & Sul (2009) to the latest available data about the regional economic performance. The Phillips & Sul's method allows us to distinguish between several most frequent 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 the economic performance at regional level. Apart from other variables, mostly inspired by the New Economic Geography and New Growth Theory and extracted from Eurostat, we concentrated on Business services.

Business services are a group of service activities ranging from legal consultancy to R&D services. These services experienced fast growth in recent years, especially in terms of employment. They are also important for their

business-to-business character and a focus on efficiency and quality enhancement. Importantly, Business services were found in previous research to play an important role in innovation and knowledge diffusion that makes them a candidate for an important growth determinant in today's knowledge based economy. At the same time, we saw an interesting ambiguity in their geographical distribution between large urban areas and the remaining regions. We therefore considered important to test their true economic impact.

The Phillips & Sul's test gave us several important insights about the regional convergence in the EU. First of all, we 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 convergence clubs. For none of these clubs we could prove an existence of the level convergence but we found that they are converging in their income growth rates instead. As far as the composition of these clubs is concerned, 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. 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.

In terms of the Central and Eastern European Countries, we did not find substantial difference between them 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, found in some previous works.

The logistic regression then showed that our main variable of interest, the Business services, has a positive impact on entering higher convergence clubs, but this impact significantly diminishes in case of the very first club. This shows that, despite the BS being beneficial for the economy, they are not a way how to enter the very top economic performance. A possible explanation for this effect could be a rather supportive character of BS for other companies'

processes and innovation, above all. They may therefore be in a position of a necessary precondition for the best club membership, in terms of an indispensable support of other key activities ensuring the top membership, without themselves guaranteeing it.

As far as the other regression results are concerned, we can see a low impact of any of the estimated variables on membership in the highest club, similarly to the Business services. Moreover, we confirmed an expected positive impact of capital cities but not of metropolitan areas in general, which supports the centre-periphery division. Notably, there is a significant positive association of membership in the higher clubs with research and patent activities. Together with the results of the innovation facilitating Business services, this possibly suggests focusing on applied science and innovation as a way to higher economic performance.

In summary, the analysis conducted in this work found a complicated pattern of regional growth convergence in the post-crisis Europe that shows a contrasting performance of regions within the member states and a significant club clustering in a metropolitan part of Western Europe. Furthermore, it confirmed the importance of research and innovation, supported, among others, by the positive impact of the specialization in Business services. This seems to support the great emphasis that is being given to the development of research institutions, and the information and knowledge economy in general, by the EU and the member states. At the same time we, however, found a lack of influence of any variable on the highest clubs. This gives space for future research.

Last but not least, it is necessary to remark that the work originally focused more directly on the CEE countries. However, in the end, our analysis concentrated on the entire Union without any special attention given to the CEEC. We therefore slightly modified the title of the thesis, compared to the proposal.

Bibliography

- ANSELIN, L. (2001): "Spatial econometrics." In B. BALTAGI (editor), "A Companion to Theoretical Econometrics," Blackwell, Oxford.
- ANSELIN, L. & S. REY (1991): "Properties of Tests for Spatial Dependence in Linear Regression Models." *Geographical analysis* **23(2)**: pp. 112–131.
- APERGIS, N., E. PANOPOULOU, & C. TSOUMAS (2010): "Old Wine in a New Bottle: Growth Convergence Dynamics in the EU." *Atlantic Economic Journal* **38(2)**: pp. 169–181.
- ARBIA, G., J. LE GALLO, & G. PIRAS (2008): "Does Evidence on Regional Economic Convergence Depend on the Estimation Strategy? Outcomes from Analysis of a Set of NUTS2 EU Regions." *Spatial economic analysis* **3(2)**: pp. 209–224.
- BALÁŽ, V. (2004): "Knowledge-Intensive Business Services in Transition Economies." *The Service Industries Journal* **24(4)**: pp. 83–100.
- BARTKOWSKA, M. & A. RIEDL (2012): "Regional Convergence Clubs in Europe: Identification and Conditioning Factors." *Economic Modelling* **29(1)**: pp. 22–31.
- BASILE, R., S. DE NARDIS, & A. GIRARDI (2001): "Regional Inequalities and Cohesion Policies in the European Union." *ISAE working papers* **23**.
- BAUMONT, C., C. ERTUR, & J. LE GALLO (2003): "Spatial Convergence Clubs and the European Regional Growth Process, 1980–1995." In B. FINGLETON (editor), "European Regional Growth," pp. 131–158. Springer.
- BENEDEK, J. & G. KOCZISZKY (2015): "Paths of Convergence and Polarization in the Visegrád Countries." In T. LANG, S. HENN, K. EHRLICH, & W. SGIBNEV (editors), "Understanding Geographies of Polarization and Peripheralization," pp. 217–234. Springer.

- BORSI, M. T. & N. METIU (2015): “The Evolution of Economic Convergence in the European Union.” *Empirical Economics* **48(2)**: pp. 657–681.
- CAMERON, A. C. & D. L. MILLER (2015): “A Practitioner’s Guide to Cluster-Robust Inference.” *Journal of Human Resources* **50(2)**: pp. 317–372.
- CARROLL, N. (2017): “*oglmx*: A Package for Estimation of Ordered Generalized Linear Models.” *CRAN, The Comprehensive R Archive Network* .
- CHAPMAN, S. A., S. COSCI, & L. MIRRA (2012): “Income Dynamics in an Enlarged Europe: the Role of Capital Regions.” *The Annals of Regional Science* **48(3)**: pp. 663–693.
- CORROCHER, N. & L. CUSMANO (2014): “The ‘KIBS engine’ of Regional Innovation Systems: Empirical Evidence from European Regions.” *Regional Studies* **48(7)**: pp. 1212–1226.
- CUARESMA, J. C., G. DOPPELHOFFER, & M. FELDKIRCHER (2014): “The Determinants of Economic Growth in European Regions.” *Regional Studies* **48(1)**: pp. 44–67.
- DALL’ERBA, S. & J. LE GALLO (2008): “Regional Convergence and the Impact of European Structural Funds over 1989–1999: A Spatial Econometric Analysis.” *Papers in Regional Science* **87(2)**: pp. 219–244.
- DESMARCHELIER, B., F. DJELLAL, & F. GALLOUJ (2013): “Knowledge Intensive Business Services and Long Term Growth.” *Structural Change and Economic Dynamics* **25**: pp. 188–205.
- ECKEY, H. F. & M. TÜRCK (2007): “Convergence of EU-Regions. A Literature Report/Convergencia de las regiones de la UE. Una revisión de la literatura.” *Investigaciones regionales* (**10**): pp. 5–32.
- FISCHER, M. M. & J. P. LESAGE (2015): “A Bayesian Space-Time Approach to Identifying and Interpreting Regional Convergence Clubs in Europe.” *Papers in Regional Science* **94(4)**: pp. 677–702.
- FRITSCH, U. & V. KUZIN (2011): “Analysing Convergence in Europe Using the Non-Linear Single Factor Model.” *Empirical Economics* **41(2)**: pp. 343–369.

- GALOR, O. (1996): “Convergence? Inferences from Theoretical Models.” *The Economic Journal* **106(437)**: pp. 1056–1069.
- GEPPERT, K. & A. STEPHAN (2008): “Regional Disparities in the European Union: Convergence and Agglomeration.” *Papers in Regional Science* **87(2)**: pp. 193–217.
- GETIS, A. & D. A. GRIFFITH (2002): “Comparative Spatial Filtering in Regression Analysis.” *Geographical analysis* **34(2)**: pp. 130–140.
- GETIS, A. & J. K. ORD (1992): “The Analysis of Spatial Association by Use of Distance Statistics.” *Geographical analysis* **24(3)**: pp. 189–206.
- GHOSH, M., A. GHOSHRAY, & I. MALKI (2013): “Regional Divergence and Club Convergence in India.” *Economic Modelling* **30**: pp. 733–742.
- GUASTELLA, G. & F. TIMPANO (2016): “Knowledge, Innovation, Agglomeration and Regional Convergence in the EU: Motivating Place-Based Regional Intervention.” *Review of Regional Research* **2(36)**: pp. 121–143.
- GUERRIERI, P. & V. MELICIANI (2005): “Technology and International Competitiveness: the Interdependence between Manufacturing and Producer Services.” *Structural Change and Economic Dynamics* **16(4)**: pp. 489–502.
- HERTOG, P. d. (2000): “Knowledge-Intensive Business Services as co-Producers of Innovation.” *International journal of innovation management* **4(04)**: pp. 491–528.
- HOSPERS, G.-J. (2002): “Beyond the Blue Banana? Structural Change in Europe’s Geo-Economy.” In “ERSA Conference Papers,” European Regional Science Association (ERSA).
- IMF (2016): “Central, Eastern, and Southeastern Europe. How to Get Back on Fast Track.” In “IMF Regional Economic Issues,” International Monetary Fund.
- KEELE, L. & D. K. PARK (2006): “Difficult Choices: an Evaluation of Heterogeneous Choice Models.” In “APSA 2004 Annual Meeting Paper,” American Political Science Association 2004 Annual Meeting.
- KOX, H. & L. RUBALCABA (2007): “The Contribution of Business Services to European Economic Growth.” In H. KOX & L. RUBALCABA (editors), “Business Services in European Economic Growth,” pp. 74–94. Springer.

- LUCAS, R. E. *et al.* (2002): “The Industrial Revolution: Past and future.” *Lectures on economic growth* pp. 109–188.
- MAASOUMI, E. & L. WANG (2008): “Economic Reform, Growth and Convergence in China.” *The Econometrics Journal* **11(1)**: pp. 128–154.
- MAGRINI, S. (2004): “Regional (di) Convergence.” *Handbook of regional and urban economics* **4**: pp. 2741–2796.
- SALA-I MARTIN, X. X. (1996): “Regional Cohesion: Evidence and Theories of Regional Growth and Convergence.” *European Economic Review* **40(6)**: pp. 1325–1352.
- MELICIANI, V. & M. SAVONA (2014): “The Determinants of Regional Specialisation in Business Services: Agglomeration Economies, Vertical Linkages and Innovation.” *Journal of Economic Geography* **15(2)**: pp. 387–416.
- MONASTIRIOTIS, V. (2011): “Regional Growth Dynamics in Central and Eastern Europe.” *LEQS Paper* (**33**).
- MONFORT, M., J. C. CUESTAS, & J. ORDONEZ (2013): “Real Convergence in Europe: A Cluster Analysis.” *Economic Modelling* **33**: pp. 689–694.
- MORA, T. (2008): “Factors Conditioning the Formation of European Regional Convergence Clubs.” *The Annals of Regional Science* **42(4)**: pp. 911–927.
- MULLER, E. & A. ZENKER (2001): “Business Services as Actors of Knowledge Transformation: The Role of KIBS in Regional and National Innovation Systems.” *Research policy* **30(9)**: pp. 1501–1516.
- PHILLIPS, P. C. & D. SUL (2007): “Transition Modeling and Econometric Convergence Tests.” *Econometrica* **75(6)**: pp. 1771–1855.
- PHILLIPS, P. C. & D. SUL (2009): “Economic Transition and Growth.” *Journal of Applied Econometrics* **24(7)**: pp. 1153–1185.
- PINHO, C., C. ANDRADE, & M. DE FÁTIMA PINHO (2010): “Regional Growth Transitions and the Evolution of Income Disparities in Europe.” *Urban Public Economics Review* (**13**): p. 66.
- PWC (2014): “Top 100 companies Europe and Germany.” Retrieved on from <https://www.pwc.de/>

- de/kapitalmarktorientierte-unternehmen/assets/
healthcare-und-pharma-werte-sind-bei-anlegern-besonders-gefragt.
pdf on 17/11/2017.
- PYLAK, K. & D. MAJEREK (2014): "Tracing Regional Economic Evolutions. KIS Impact on Growth." *Procedia Economics and Finance* **14**: pp. 515–524.
- RODRÍGUEZ-BENAVIDES, D., F. LÓPEZ-HERRERA, & F. VENEGAS-MARTÍNEZ (2014): "Are there Economic Convergence Clubs in Latin America?" *Journal of Economics* **2(3)**: pp. 113–123.
- SHEARMUR, R. & D. DOLOREUX (2009): "Place, Space and Distance: Towards a Geography of Knowledge-Intensive Business Services Innovation." *Industry and Innovation* **16(1)**: pp. 79–102.
- SMĘTKOWSKI, M. (2013): "Regional Disparities in Central and Eastern European Countries: Trends, Drivers and Prospects." *Europe-Asia Studies* **65(8)**: pp. 1529–1554.
- SMETKOWSKI, M. & P. WÓJCIK (2012): "Regional Convergence in Central and Eastern European countries: A Multidimensional Approach." *European Planning Studies* **20(6)**: pp. 923–939.
- SZENDI, D. (2013): "The Convergence Process and the Effects of the Economic Crisis in Central-Eastern Europe." *Romanian Review of Regional Studies* **9(1)**.
- TOIVONEN, M. *et al.* (2004): *Expertise as Business: Long-Term Development and Future Prospects of Knowledge-Intensive Business Services (KIBS)*. Ph.D. thesis, Helsinki University of Technology.
- WILLIAMS, R. (2009): "Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients across Groups." *Sociological Methods & Research* **37(4)**: pp. 531–559.
- ZAROTIADIS, G. & A. GKAGKA (2013): "European Union: a Diverging Union?" *Journal of Post Keynesian Economics* **35(4)**: pp. 537–568.

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), Łódzkie, 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, An-

dalucía, Dél-Dunántúl, Észak-Magyarország, Észak-Alföld, Dél-Alföld,
Campania, Puglia, Calabria, Sicilia, Podkarpackie, Warminsko-Mazurskie,
Norte, Nord-Est, West Wales and The Valleys

Appendix B

Relative Transition Paths

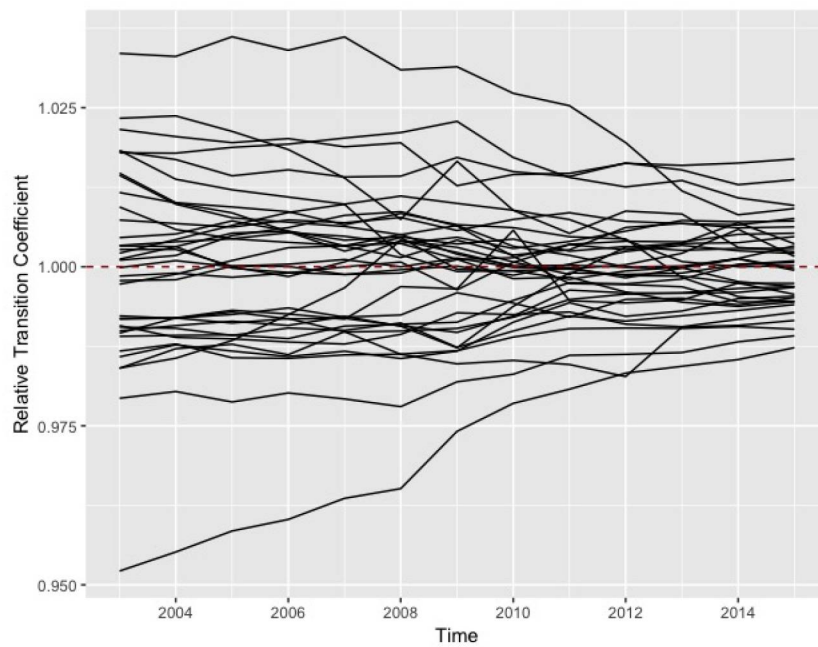


Figure B.1: Transition paths for members of Club 2 (merged)

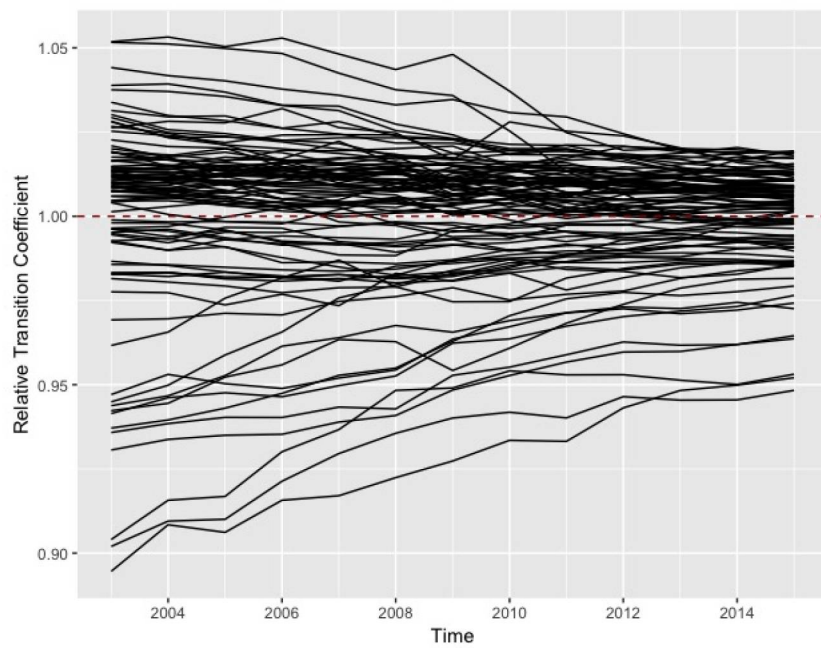


Figure B.2: Transition paths for members of Club 3 (merged)

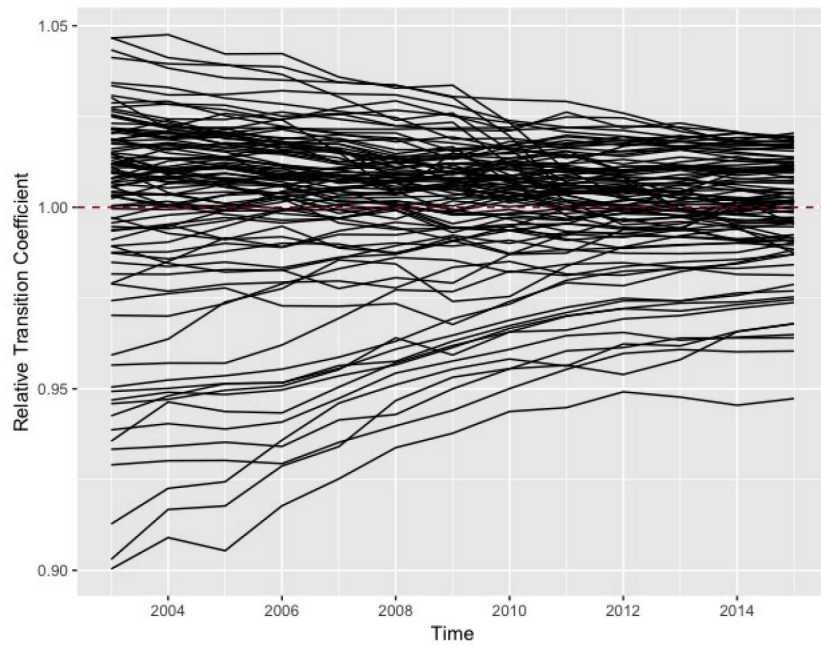


Figure B.3: Transition paths for members of Club 4 (merged)

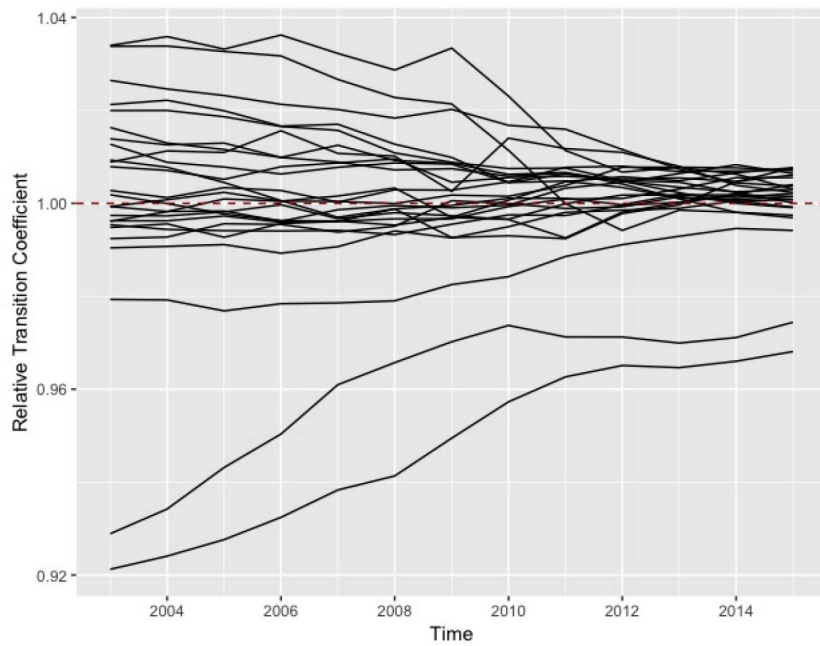


Figure B.4: Transition paths for members of Club 3 before merging

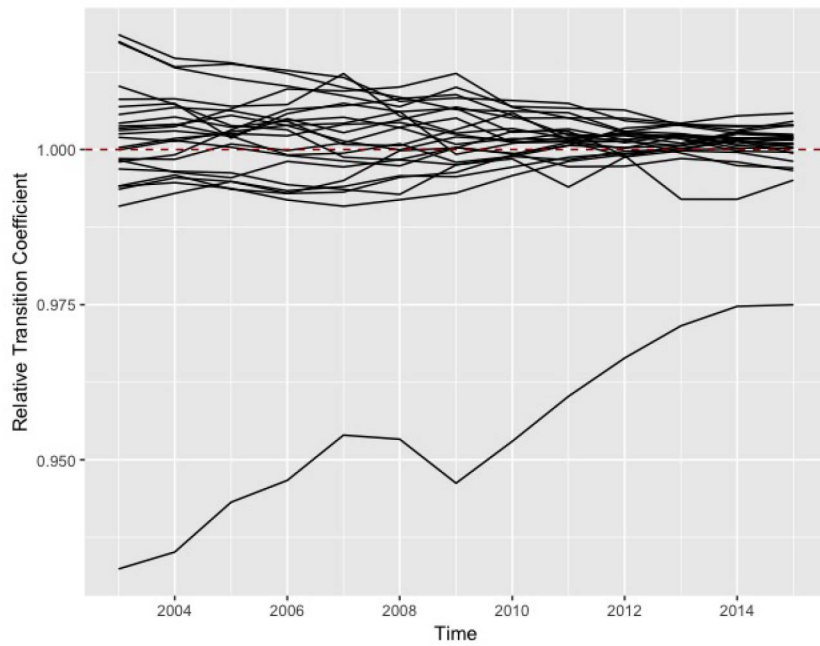


Figure B.5: Transition paths for members of Club 4 before merging

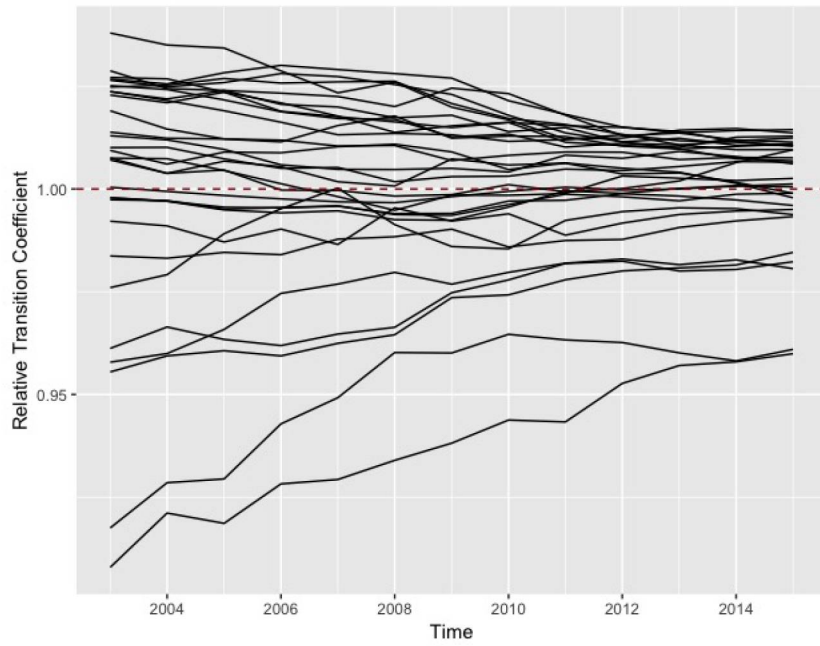


Figure B.6: Transition paths for members of Club 5 before merging

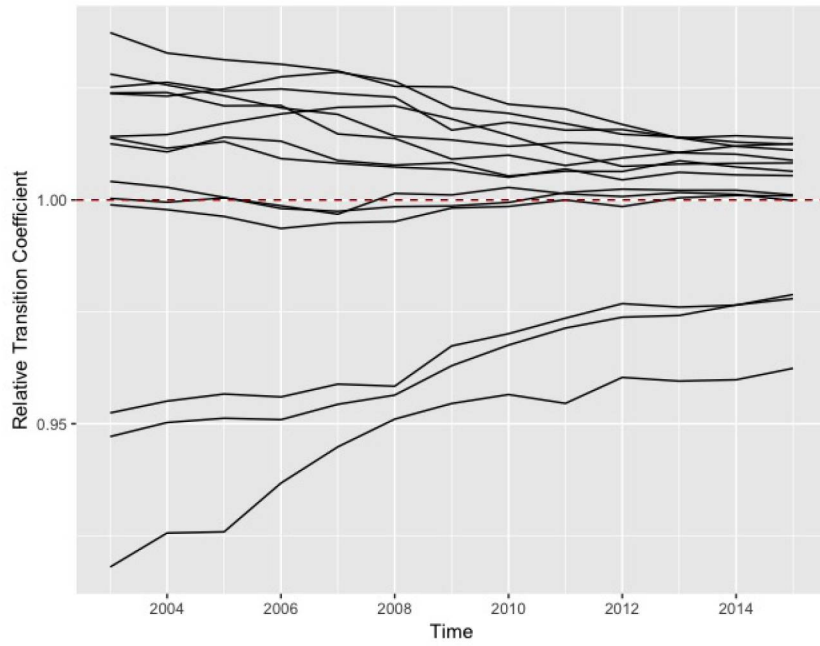


Figure B.7: Transition paths for members of Club 6 before merging

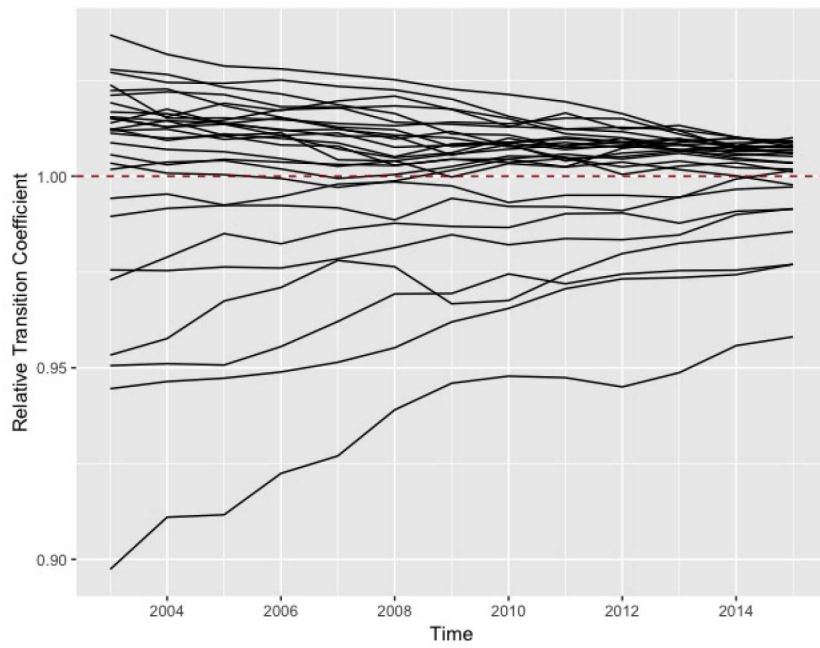


Figure B.8: Transition paths for members of Club 7 before merging

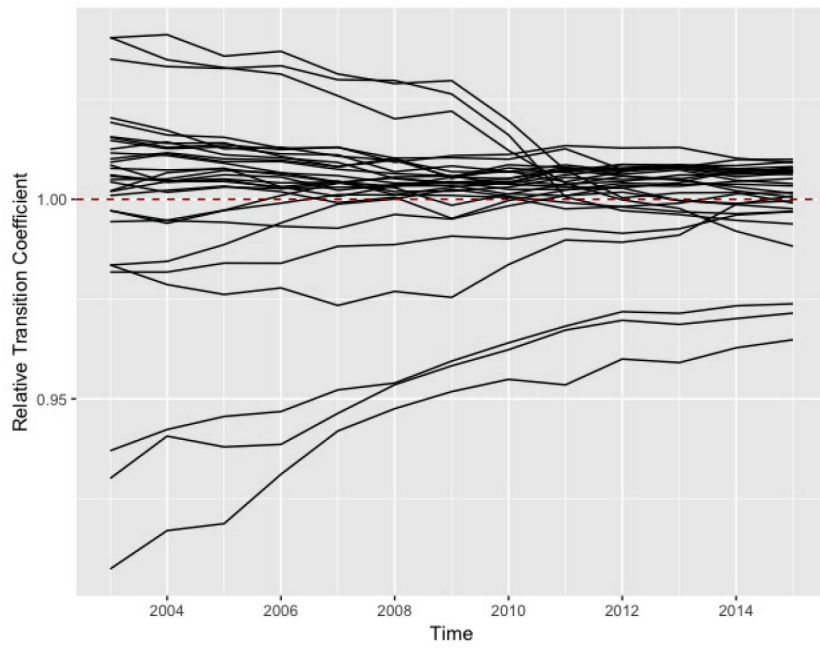


Figure B.9: Transition paths for members of Club 8 before merging

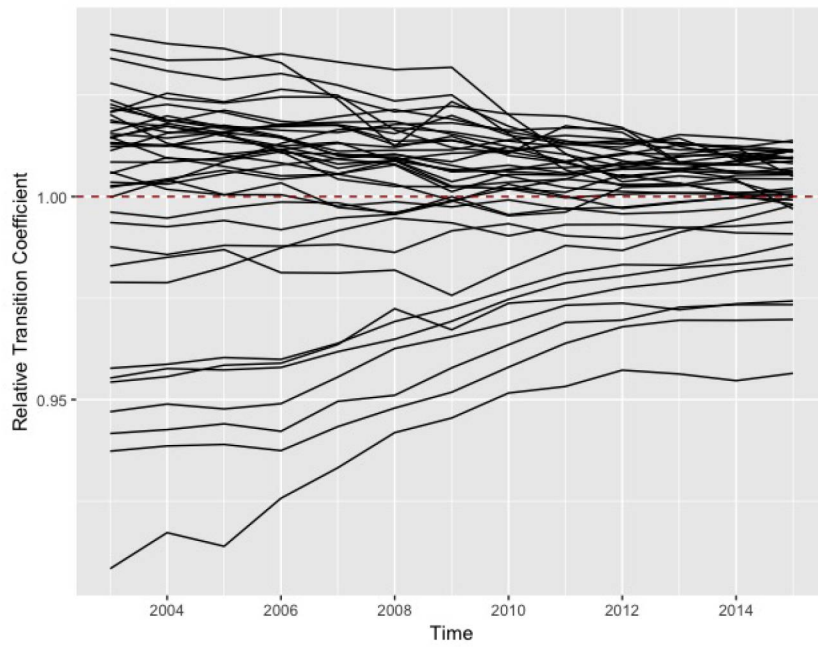


Figure B.10: Transition paths for members of Club 9 before merging

Appendix C

Regression Outputs

Table C.1: Regression results: Logistic regression with Convergence Clubs before merging and FAMD imputation

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

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

Table C.2: Marginal effects: Logistic regression with Convergence
Clubs before merging and FAMD imputation

| | Club 10 | Club 9 | Club 8 | Club 7 | Club 6 |
|---------------------|-----------|-----------|-----------|-----------|-----------|
| Population growth | -0.114*** | -0.353*** | -0.424*** | -0.386*** | -0.137** |
| Population density | 0.000*** | 0.000*** | 0.000*** | 0.000*** | 0.000* |
| Young\old ratio | 0.029** | 0.091** | 0.109** | 0.100** | 0.035* |
| Tertiary edu. share | -0.012 | -0.037 | -0.044 | -0.040 | -0.014 |
| Share of scientists | -3.569*** | -11.10*** | -13.31*** | -12.13*** | -4.292*** |
| Business Services | -0.545*** | -1.695*** | -2.032*** | -1.852*** | -0.655*** |
| Infrastructure | -0.078*** | -0.243*** | -0.292*** | -0.266*** | -0.094** |
| Wage difference | -0.001 | -0.003 | -0.004 | -0.004 | -0.001 |
| Internet use | 0.000 | -0.001 | -0.001 | -0.001 | 0.000 |
| Old EU states | 0.019 | 0.060** | 0.077** | 0.082 | 0.039** |
| Spec. agriculture | -0.044*** | -0.136*** | -0.163** | -0.149 | -0.053** |
| Spec. industry | -0.091*** | -0.284*** | -0.340** | -0.310** | -0.110** |
| Metrop. regions | 0.009 | 0.029 | 0.036 | 0.037 | 0.015 |
| Capitals | -0.021** | -0.066*** | -0.088*** | -0.098*** | -0.051** |
| Top universities | 0.031* | 0.087* | 0.084* | 0.054* | 0.007 |
| Corporations | -0.006 | -0.019 | -0.024 | -0.023 | -0.009 |
| Patent activity | 0.000*** | -0.001*** | -0.001*** | -0.001*** | 0.000** |
| Connectivity | -0.029 | -0.089 | -0.107 | -0.098 | -0.035 |
| | Club 5 | Club 4 | Club 3 | Club 2 | Club 1 |
| Population growth | 0.115 | 0.414*** | 0.479*** | 0.379*** | 0.026** |
| Population density | 0.000 | 0.000*** | 0.000*** | 0.000*** | 0.000** |
| Young\old ratio | -0.030 | -0.107** | -0.124** | -0.098** | -0.007 |
| Tertiary edu. share | 0.012 | 0.043 | 0.050 | 0.039 | 0.003 |
| Share of scientists | 3.604 | 13.01*** | 15.05*** | 11.90*** | 0.828** |
| Business Services | 0.550 | 1.987*** | 2.298*** | 1.818*** | 0.126** |
| Infrastructure | 0.079 | 0.285*** | 0.330*** | 0.261*** | 0.018** |
| Wage difference | 0.001 | 0.004 | 0.005 | 0.004 | 0.000 |
| Internet use | 0.000 | 0.001 | 0.001 | 0.001 | 0.000 |
| Old EU states | 0.021 | -0.066** | -0.112** | -0.112* | -0.009 |
| Spec. agriculture | 0.044 | 0.160*** | 0.184*** | 0.146*** | 0.010** |
| Spec. industry | 0.092 | 0.333*** | 0.385*** | 0.304*** | 0.021** |
| Metrop. regions | 0.000 | -0.035 | -0.047 | -0.041 | -0.003 |
| Capitals | -0.050 | 0.062*** | 0.138*** | 0.162** | 0.013* |
| Top universities | -0.056 | -0.077* | -0.074* | -0.053* | -0.004* |
| Corporations | 0.002 | 0.023 | 0.029 | 0.025 | 0.002 |
| Patent activity | 0.000 | 0.001*** | 0.001*** | 0.001*** | 0.000** |
| Connectivity | 0.029 | 0.105 | 0.121 | 0.096 | 0.007 |

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

Table C.3: Regression results: Logistic Regression with merged Convergence Clubs and kNN imputation

| <i>Dependent variable:</i> Convergence Club | | | |
|---|-----------------------|-------------------|---------------------|
| Population growth | 6.446*** (0.970) | Old EU states | -1.104** (0.531) |
| Population density | 0.001** (0.0003) | Spec. agriculture | 3.935*** (0.409) |
| Young\old ratio | -1.146* (0.692) | Spec. industry | 2.786*** (0.193) |
| Tertiary edu. share | 1.000 (0.864) | Metrop. regions | -0.442 (0.433) |
| Share of scientists | 186.879*** (0.016) | Capitals | 1.438** (0.575) |
| Business Services | 22.159*** (0.018) | Top universities | -0.987 (0.654) |
| Infrastructure | 4.144*** (0.673) | Corporations | 0.268 (0.418) |
| Wage difference | 0.024 (0.041) | Patent activity | 0.011*** (0.003) |
| Internet use | 0.013 (0.025) | Connectivity | 2.186 (1.387) |

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

Table C.4: Marginal Effects: Logistic Regression with merged Convergence Clubs and kNN imputation

| | Club 5 | Club 4 | Club 3 | Club 2 | Club 1 |
|---------------------|-----------|-----------|----------|----------|---------|
| Population growth | -0.139*** | -1.275*** | 0.934*** | 0.447*** | 0.032** |
| Population density | 0.000** | 0.000*** | 0.000*** | 0.000** | 0.000* |
| Young\old ratio | 0.025* | 0.227* | -0.166 | -0.080* | -0.006 |
| Tertiary edu. share | -0.022 | -0.198 | 0.145 | 0.069 | 0.005 |
| Share of scientists | -4.036*** | -36.95*** | 27.08*** | 12.97*** | 0.936** |
| Business Services | -0.479*** | -4.381*** | 3.211*** | 1.538*** | 0.111** |
| Infrastructure | -0.089*** | -0.819*** | 0.600*** | 0.288*** | 0.021** |
| Wage difference | -0.001 | -0.005 | 0.003 | 0.002 | 0.000 |
| Internet use | 0.000 | -0.002 | 0.002 | 0.001 | 0.000 |
| Old EU states | 0.019 | 0.193** | -0.104** | -0.099 | -0.008 |
| Spec. agriculture | -0.085*** | -0.778*** | 0.570*** | 0.273*** | 0.020** |
| Spec. industry | -0.060*** | -0.551*** | 0.404*** | 0.193*** | 0.014** |
| Metrop. regions | 0.008 | 0.083 | -0.055 | -0.035 | -0.003 |
| Capitals | -0.021** | -0.232*** | 0.092** | 0.148** | 0.013 |
| Top universities | 0.030 | 0.203 | -0.176 | -0.053 | -0.004 |
| Corporations | -0.005 | -0.051 | 0.035 | 0.020 | 0.001 |
| Patent activity | 0.000*** | -0.002*** | 0.002*** | 0.001*** | 0.000** |
| Connectivity | -0.047 | -0.432 | 0.317 | 0.152 | 0.011 |

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

Table C.5: Regression results: Weighted Logistic Regression with merged Convergence Clubs and kNN imputation

| <i>Dependent variable:</i> Convergence 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. industry | -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 | | |
| <i>Significance levels:</i> | *p<0.1; **p<0.05; ***p<0.01 | | |

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

| <i>Dependent variable:</i> Convergence Club | | | |
|---|-----------------------------|-------------------|---------------------|
| Population growth | 6.805*** (0.949) | Old EU states | -0.691 (0.544) |
| Population density | 0.001** (0.0003) | Spec. agriculture | 3.810*** (0.400) |
| Young\old ratio | -1.049* (0.630) | Spec. industry | 3.744*** (0.203) |
| Tertiary edu. share | 0.298 (0.838) | Metrop. regions | -0.494 (0.432) |
| Share of scientist | 191.189*** (0.017) | Capitals | 1.535*** (0.569) |
| Business Services | 25.876*** (0.025) | Top universities | -1.053 (0.670) |
| Infrastructure | 3.400*** (0.666) | Corporations | 0.345 (0.390) |
| Wage difference | 0.022 (0.042) | Patent activity | 0.010*** (0.002) |
| Internet use | 0.013 (0.024) | Connectivity | 2.133 (1.362) |
| Observations | 244 | | |
| <i>Significance levels:</i> | *p<0.1; **p<0.05; ***p<0.01 | | |