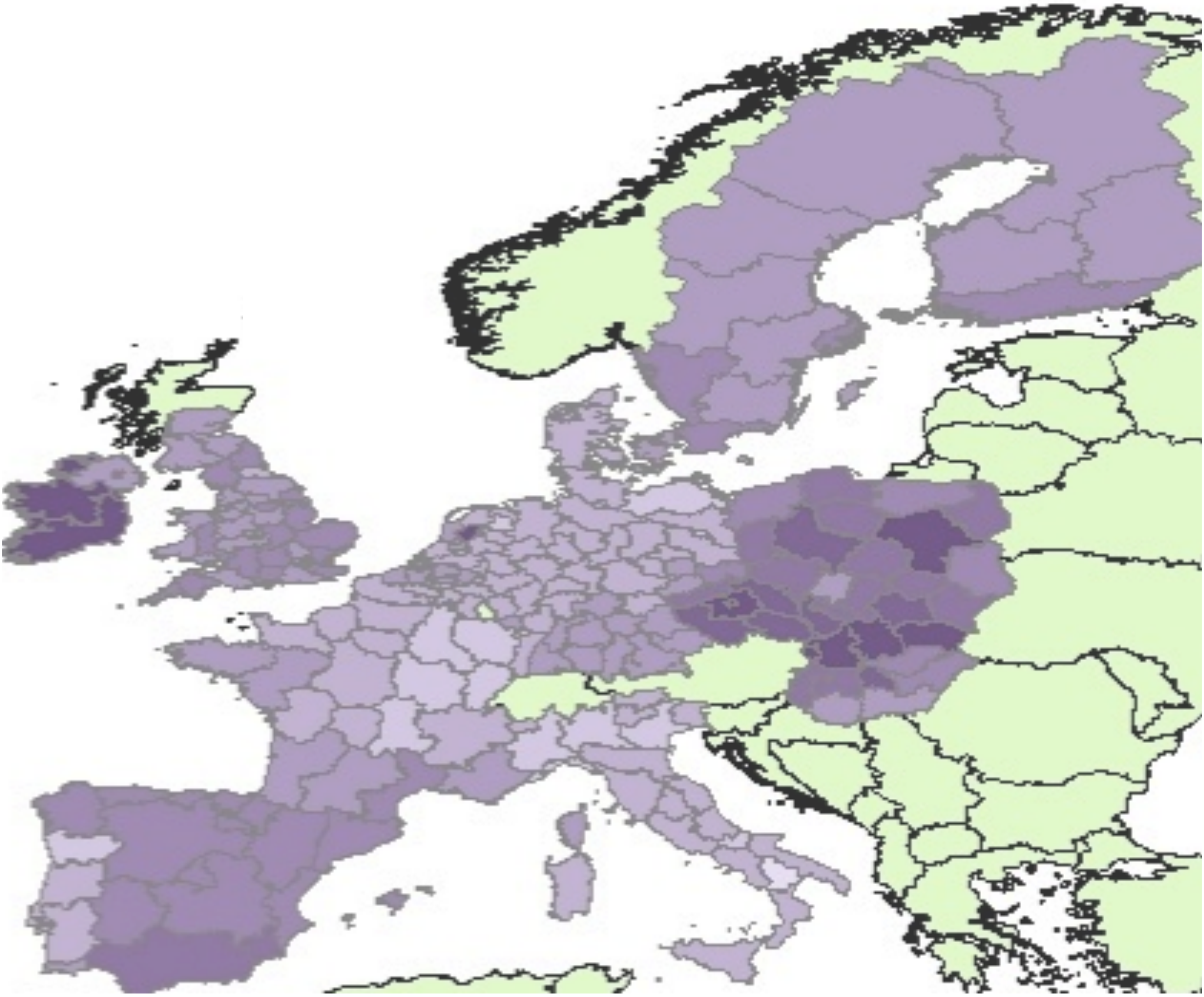
Applied Economics Research Course, Economic Geography, Utrecht University

Regional Economic Growth in Europe

Analysis of Regional Dependence and Structure

Stefan de Geus
Supervisor: Prof. Dr. F. G. van Oort



Regional Economic Growth in Europe

Analysis of Regional Dependence and Structure

Master Thesis Economic Geography

Utrecht University

Faculty of Geosciences

Stefan de Geus

Supervisor:

- Prof. Dr. F. G. van Oort (Utrecht University)

February 2013

Acknowledgements

This thesis is the product of just over six months of theoretical and empirical research. It all started with Frank van Oort, who made me interested in the topic with the dataset he used in the Masters course Spatial Economic Policy in Europe. When Frank van Oort offered me the opportunity to use this dataset to study one of the most fascinating topics in economic geography, namely regional economic growth differences, including spatial dependence, I did not doubt for a single second to accept the challenge. With my econometric background and my geographical interests, it was the perfect topic for my Master thesis. I want to thank Frank van Oort, who supported me from the very beginning with helpful comments and suggestions and made me feel more like a colleague, rather than a student.

I also want to thank Lieselotte Hoegee for helping me with the style and grammar and Maarten van Smeden, who gave me valuable feedback and insights throughout the empirical research. He especially had strong influence on the models used in the empirical research and the different tests to evaluate the dataset. Of course, this research would not have been possible without the large and valuable databases of the Netherlands Environmental Assessment Agency and Bureau van Dijk.

Finally, I want to thank both my girlfriend for her support and patience during this research project and my parents for accommodating excellent workspace in Utrecht as well as in Bielefeld.

Stefan de Geus
Utrecht, February 22nd 2013

Abstract

The economic structure of European regions is said to be of significant influence on economic growth. To get a more detailed view on the influence of this economic structure the related variety, unrelated variety and specialization of the regions are examined for their influence on economic growth. Several country level analyses researched this issue up to now. For the first time, in this thesis, a EU-wide analysis is done taking into account the spatial dependence of the regions, which is included due to the heterogeneity of economic performance in European regions. This heterogeneity is also captured in regime-tests where different groups of regions are tested and compared according to specific regimes.

Although economic performance is distributed heterogeneously between European regions, the regions are converging in their performance. In search for the best policy it is found that a region specific policy mix is best to achieve maximum growth. Investing in human capital is found to be a good choice when the region lags behind in productivity or when the region has high unemployment. Related variety is found to be of positive influence on employment growth while unrelated variety is to a lesser extent positive for employment growth. Specialization is positively related with productivity growth and of negative influence on unemployment.

Table of Contents

Introduction

Regional economic growth, how can it be reached? What is the driving force behind growth and why are there such great differences in regional economic performance across the world? This is one of the most interesting topics in geographical economics nowadays. When considering Europe, one can see that the European Union tries to reduce these differences using convergence policy, which was first recorded officially at the Treaty of Maastricht in 1992, where five goals were established to unify Europe. Nowadays the European Union succeeds in reducing economic performance differences on a national scale, but regional differences are still considerable (Martin & Sunley, 1998; Puga, 2002; Frenken & Hoekman, 2006; Cornett & Sorensen, 2010).

It is a difficult task for the EU to discover the forces behind regional growth. The opening of national borders in Eastern Europe, the increased internal market of the EU and the fast technological and scientific advancements affected regional development over the last decades. Some regions suddenly faced an extremely competitive open market. Due to this course of events regions should not be seen as administrative sub-divisions of countries, but as integral parts of the EU economic space. Therefore, regional analysis has proven to be much more important than initially expected (Petrakos et al, 2007).

It should be noted that countries with greater economic decentralization, like Germany and the US, were more successful in history, which is partly the result of the decentralized character of these countries. Due to centralization of political power at the national level, there is urban priority and a size distribution of cities favoring large cities, which definitely contributes to heterogeneity and divergence (Porter, 2003 and Kim & Law, 2012). In order to induce more growth, economic policy should thus particularly focus on the regional level.

This thesis will focus on regional economic performance in Europe. Three different components of economic performance will be analyzed: employment growth, productivity growth and unemployment growth. By using a large dataset including just over 200 regions, this thesis tries to demonstrate what influences the components for economic growth in European regions. Special attention will be paid to (1) the structure of the regions and (2) the spatial dependence of the regions. According to literature these two concepts have an extensive influence on regional economic growth and up till now only few studies have looked into these concepts in relation to regional economic growth at European scale. Questions that arise from literature regarding these concepts are the following: Does the region have a diverse production sector or is there a focus on one specific product or sector? In case of a diverse production sector, can one speak of related variety or unrelated variety? Is the agglomeration of economic activities positive or negative for economic growth? And what is the influence of the spatial location of the region? In other words, what is the influence of economic interaction with neighboring regions? In particular this last concept of economic interaction is underexposed in today's literature. This thesis wants to give an answer to all of these questions, with the following central research question as its main guide.

What is the influence of related variety, unrelated variety and specialization on the economic performance of European regions taking into account their spatial location and characteristics and how can these processes be translated into European-wide policy?

Relevance

Regional policy in the European Union nowadays focuses on smart specialization and place-based development. Smart specialization is a strategy where supporting research and innovation is the key to economic growth. With smart specialization the knowledge potential of the region is maximized, regardless of whether this is a high-tech or low-tech region. Smart specialization though is a concept in sector-growth literature, which should not be applied to regional policy without any adaptations, because specialization works not only in smart sectors (McCann & Ortega-Argiles, 2011).

Place-based development is a strategic approach to economic development according to which regions must plan and develop the local economy in response to possibilities and limitations of that particular region. Policies are based on the place or region itself. The region focuses on few sectors and specializes in these sectors. It is assumed that the region knows which sectors to focus on (Barca et al, 2012). Barca (2009) says that place-based policies are the best way to tackle the "persistent underutilization of potential and reducing persistent social exclusion" (Barca, 2009, p. VII) in all areas of Europe. Political focus is mainly directed to the most successful sectors, which can result in less variety in the region. According to evolutionary economic geography, policy should focus on the different types of variety in order to attain regional economic growth. All approaches bring forward various insights to the regional structure and economic performance. This thesis considers these several approaches and builds on earlier studies regarding the relationship between variety, specialization and economic growth (Witte, 2011 and Frenken et al, 2007). Over the last two decades, much research has been done on diversity and specialization. Nevertheless, thus far the specialization-diversity debate did not lead to any conclusive results. Therefore, focus should shift from this debate to the transfer mechanisms of knowledge and knowledge spillovers. The models used in previous specialization-diversity studies did not take these flows of knowledge and spatial dependence into account (Van Oort, 2013).

This thesis adds to literature by considering diverse production sectors in terms of their related variety or unrelated variety. Furthermore, this empirical research is not just based on a single country, like has been done in most previous researches, but on a large number of countries in Europe. This Europe-wide approach taking into account related and unrelated variety is unique and especially interesting for theory and policy implications on a European scale. In addition, this research considers the influence of spatial dependence, which makes the outcomes more reliable compared to earlier studies, since spatial dependence between regions in Europe is undeniably present. According to the literature discussed later in this thesis, there is heterogeneity between the European regions. Neglecting the existence of spatial dependence and its influence on economic development may result in incorrect conclusions, as will be argued later.

Contents

The first part of this thesis will address the existing literature, which deals with economic growth, specifically focusing on sector structure and spatial dependence. The second part summarizes the different theories resulting in a research model with several hypotheses and an accessory conceptual model. It also describes the dataset and methodology used for the empirical research. In part three the results of this empirical research are presented. The fourth chapter analyzes and interprets the outcomes of the empirical research, followed by an overall conclusion that will try to give an answer to the central research question.

Important Concepts

Related Variety and Unrelated Variety

Variety is a concept mainly used in mathematics to describe dispersions. The more variety the larger is the dispersion. In economics, variety can be used as product variety, which refers to the number of variants within a specific product group (Lancaster, 1990). A region where firms make all kinds of different products is labeled a region with high variety, while a region with for example mainly automobile factories is a region with low variety. In this thesis variety is split into related and unrelated variety. Related variety means that the variety is closely related, like for example a car manufacturer and a tire manufacturer. They are operating in the same sector, but they are making different products. Unrelated variety is quite different; firms are operating in different industries making different products. An example of unrelated variety is a hospital and a bakery.

Spatial Analysis and Spillover Effects

Spatial analysis is a method of analysis that accounts for spatial dependence. An example of spatial analysis is the determination of house values by James LeSage. The idea is that areas with high property values might be adjacent to other high value areas, which results in a spatial trend in the outcome variable. The example shows that independent variables can change in magnitude or even become insignificant when accounting for spatial dependence. The spatial analysis also improves the model fit (LeSage & Pace, 2009).

In spatial analysis, the nested structure of a dataset is taken into account. When analyzing European regions without taking into account the spatial dependence, the interconnectedness between the regions is ignored; all regions are seen as independent cases. This is likely to be problematic for the model, due to the expected spatial dependence of the geographically connected regions. Standard estimation techniques can result in biased estimates in an OLS model. Therefore, when using a dataset containing over 200 regions in Europe with border connections, it is better to do spatial analysis instead of standard estimation techniques. Further explanation regarding this type of analysis is done in the methodology section later in the thesis.

In spatial analysis, economic interaction is considered by looking at the spillover effects from region to region. Spillover effects are side effects of an activity or process that affect those not directly involved. Spillovers can be negative or positive. Pollution by a car for example has a negative spillover effect, as the driver enjoys his ride while others not enjoying the ride may be affected by the pollution. An example of a positive spillover effect is the beautiful garden of your neighbor. While your neighbor pays for the garden and enjoys it, so will you without paying anything for it. In this thesis only positive spillover effects will be considered.

Cumulative causation

Cumulative causation assumes that knowledge is immobile. Cumulative causation means that multiple, successive changes are set in motion by one single event. This chain reaction can be positive or negative depending on the "single event" that occurred. The establishment of new firms in a region results for example in positive causation, while closing of firms has the opposite effect. The chain effect is said to be cumulative which means that effects are getting more positive or negative over time. Extremely important are the initial conditions of the region and without balancing policies divergence is the most possible outcome regarding the spread in regional economic growth. Actually, three stages are described in cumulative causation theory. The first is pre-industrial where regional inequalities are small. The second is the phase where the cumulative causation effect is working and divergence occurs. And a third and last stage is where the so-called spread effect stimulates growth in periphery and neighbor regions. Interregional interactions are related to regional economic development (Myrdal, 1957; Kaldor, 1970). As the pre-industrial phase is over in European regions, the second and third phases are most important in this thesis. The question is whether there is a cumulative causation effect visible or a spread effect.

1. Literature Review

This literature overview starts with an introduction to economic geography. For specialists in this field this is facultative reading, but it might be useful for a layman to grasp the subject matter a little better. After this introduction, theories are discussed which deal with economic growth, specialization, related variety, unrelated variety and spatial dependence. This literature review ends with the description of several regimes that may affect the total model of economic growth due to heterogeneity of the European regions.

1.1 Economic Geography

Economic geography is, as the name itself states, a combination of economics and geography. It has to do with the location, distribution and spatial organization of economic activities. It deals with a lot of subject matters, for example globalization, international trade, real estate, transportation and economics of agglomeration. Economic geography is a very broad discipline with a number of methodologies and approaches. This thesis, due to its regional focus, considers economic geography on a regional scale, where it looks at economic regionalization and local economic development.

Unfortunately, a single generalized and widely accepted theory on regional economic growth is missing in today's available literature. Nevertheless, a summary of the various theories at hand can give a good insight on what factors are expected to have an influence on economic growth. The most popular theories for regional economic growth are the neoclassical growth theory, the new growth theory, the new economic geography theory, the evolutionary economic geography theory, the urbanization theory and the agglomeration theory. Most of these theories are not specifically focused on regional growth, but they might give an indication of what is important for growth in general. In the next part these theories will be discussed to understand what can foster regional economic growth.

1.2 From the Neoclassical Growth Theory to the Evolutionary Economic Geography

One of the first developed theories in economic geography is the theory of the location of industries by Alfred Weber (Weber, 1929). The theory is based on Von Thünens model, which states that the price of products is not only dependent on land prices, but also on transportation costs (Von Thünen, 1966). Weber's theory helped finding optimal locations for manufacturing plants with minimal costs. This theory was very relevant due to the industrial revolution and development of several transportation options at that time. Finding the optimal location has to do with three factors; transportation costs, cost of labor and the concentration of firms in the area. The last factor can both have a positive as well as a negative effect. Concentration of activities lowers transportation costs to suppliers and customers, but overconcentration or a concentration of the wrong type of industry can result in negative effects.

1.2.1 Neoclassical Growth Theory

Weber's theory of the location of industries belongs to the neoclassical location theories, which focus particularly on industrial location. Out of these neoclassical thoughts the neoclassical growth theory arose. The neoclassical growth theory shows how steady economic growth is obtained on the basis of productivity, capital, population growth and technology. Labor and capital should be at an optimal level, because together they maximize economic growth. When new technology comes in, the amount of labor and capital needs to be adjusted to a new optimal level in order to reach new growth equilibrium. The main concepts in neoclassical growth theory are constant returns to scale, diminishing marginal productivity of capital, substitutability between capital and labor and exogenously determined technical progress.

Constant returns to scale means that the scale of production is at the exact level so that input equals output. Firms always try to come to this optimal level, as they always try to get as much output for what they put in. As long as the output exceeds the input, a firm will raise input till constant returns to scale is reached.
Diminishing marginal productivity of capital means that the impact of extra capital invested is always less than the impact of capital invested earlier. Capital and labor are substitutable and at some point an optimal amount of capital and labor is reached. Technological progress can result in more growth, but this is assumed to be exogenous according to the neoclassical growth theory (Solow, 1956). Solow's growth model, where diminishing returns is one of the key factors, gives rise to convergence. Developing countries tend to grow faster than developed countries, because these diminishing returns are higher in poorer regions than in rich regions. Also poorer regions can benefit by replicating technological knowledge and institutions of richer countries. In some European regions the steady state, where constant returns to scale are reached, is not reached yet, which can result in even more convergence. Knowledge is assumed to be perfectly mobile according to neoclassical growth theory, which is hard to believe nowadays.

1.2.2 New Growth Theory

Some 30 years after Solow's neoclassical growth theory the endogenous growth theory was developed, which states that growth can also be generated endogenous instead of exogenous as was stated by the neoclassical growth theorists. The endogenous growth theory shows that productivity growth can be obtained with innovation and investment in human capital. An innovative, knowledge driven economy can generate positive externalities and spillover effects, which drives economic growth from the inside (Romer 1986).
This theory by Romer, which showed how growth could also be endogenous, resulted in the new growth theory. Human desires and unlimited wants, which are endogenous by character, are considered to be the main drivers of productivity growth and thus economic growth. Instead of increasing labor or capital, new knowledge results in economic growth. Knowledge is not assumed to be perfectly mobile anymore like in neoclassical growth theory and one of the interesting properties of knowledge is that once it is there, using the knowledge has zero marginal costs; producing an extra unit using the knowledge does not cost any extra money once the knowledge is available. All types of knowledge contribute to growth and development through knowledge may even provide firms with a monopoly position, due to their advantages. The technological development, which results from the knowledge, makes sectors diverse in production. Every firm has different knowledge and therefore its own ways to come to the same product. Knowledge is the main driver of productivity growth according to the new growth theory. Economies with larger stocks of human capital will experience faster growth (Romer, 1990). The result of the uneven spread of knowledge is that there is also uneven spread in economic activity. Knowledge is geographically bounded and knowledge is build into routines of individuals and organizations. Development will therefore be more and more dependent on path dependence (Cortright, 2001).
Path dependence implies that decisions are limited by decisions in the past, even when the circumstances are not longer relevant. A consequence of this path dependence and the resulting uneven spread of knowledge and economic activity is that economies can experience "lock-in" to particular technologies and/or locations.
The most well known example of lock-in is the example of the QWERTY-keyboard. The technology that made typewriters use QWERTY classification of characters resulted in a lock-in for computers nowadays. As everybody has the knowledge to use this type of keyboard there is a dependence on this type of keyboard. While other classifications may be more efficient we are "locked in" to this type of keyboard as a result of technologies in the past (David, 1985).
Due to the path dependence and the possibility of lock-in, economic growth through knowledge accumulation can be bounded. Compare this path dependence with the earlier described concept of cumulative causation.

Economic performance is, in both concepts, dependent on the initial conditions of the regions. Also important is the phase in each concept in which the region operates. Regions that are developed less and operate in phase 2 of the cumulative causation concept don't have to deal with lock-in or spread effects and these regions can grow in a different way compared to regions in a later phase, who need to grow through interregional interactions.

1.2.3 New Economic Geography

In the same period new economic geography evolved, introduced by Paul Krugman. This theory is based on the new trade theory in which Krugman explained that the former trade theory based on comparative advantages no longer holds. A country or region had a comparative advantage to another country when it produced a good at lower opportunity costs than other countries or regions. According to new trade theory advantages are now enjoyed from diversity and expansion, because consumers prefer diversity in brands and production favors economies of scale (Krugman, 1979).
Economies of scale are cost advantages resulting from expansion. As firms become larger they benefit from these economies of scale. The average cost per unit produced decreases as output increases. Take for example a car factory that finishes two cars a day; producing an extra car makes the average cost per car produced lower as the factory can use the same machines for the extra car.
This new trade theory evolved into new economic geography. The same "economies of scale" principle also holds for regional economic geography. Regions benefit from more economic activity in the region through economies of scale. With higher concentration of production, regions will be more profitable. Firms tend to relocate to profitable regions and therefore economic activity concentrates in a few, densely populated places, which brings higher levels of income (Krugman, 1991a). Economic activity and economic growth is geographically unbalanced, as some regions experience high growth and others are suffering. The uneven spread may lead to the same lock-in as in path dependence.

1.2.4 Evolutionary Economic Geography

The theory of Evolutionary Economic Geography (EEG) makes use of many of the earlier discussed theories and tries to provide more insight into the process of economic development. It builds on the earlier discussed new growth theory and can also be seen as economic geography with an evolutionary approach. Economic geography is always subject to the past; the new economic geography shows this with the economies of scale and the new growth theory shows this with the concept of path dependency. Adding an evolutionary approach to economic geography thus seems reasonable according to the latest most important theories on economic growth. Technology is still thought to be the main driver of economic growth, but in EEG technological growth is path dependent and subject to history and its routines.
Path dependence, which implies that decisions are influenced by decisions made in the past, has positive and negative effects on regional economic performance. The positive effect has to do with the innovative milieu. Firms make decisions to locate closely to related firms due to the huge advantages in supply of human capital, knowledge, information linkages, network externalities and supportive institutions (Boschma & Lambooy, 1999). A region like Silicon Valley is a typical example of a region enjoying positive path dependence effects, because due to the abundance of firms in the technological sector even more firms are locating in Silicon Valley, which means growth of the economic activity in the region. The decision of new firms to locate in Silicon Valley is based on other firms' earlier decision to locate in that certain area, which is explained as path dependence. The negative effect has to do with the earlier discussed "lock-in" effect. Knowledge is build into routines of individuals and organizations. These routines are hard to copy and largely determine the competitiveness of a firm. Routine replication is the process that is behind routine distribution. Routine replication can be done within a firm, through spin-offs from parent-firms and through labour mobility (Frenken & Boschma 2007).

Most of this replication takes place locally. This is of course the case for replication within firms, but also for spin-offs which tend to locate closely to the parent firm, and for employees that change jobs, because they usually find a new job in the same labor market area (Frenken & Boschma 2007).
Besides path dependence and routine replication two other important issues in EEG are economies of scope and Jacobs externalities. Economies of scope imply advantages due to variety at the level of the firm. Firms gain advantages when they make two or more different products, which are related to each other. It is almost the same as economies of scale, but the difference is that in this case it is not simply about producing more of the same product, it is now about producing similar goods. A company that produces TV's can gain advantages when producing more TV's (economies of scale) but it can also gain advantages when making displays for computers (economies of scope).
Jacobs externalities are advantages at a larger scale. Named after Jane Jacobs, these externalities imply advantages gained from variety at the urban level. The more variety, the more external spillovers can be enjoyed. External spillovers imply that ideas are exchanged among individuals outside the firm, where internal spillovers imply exchange of ideas inside the firm. These two issues both work through the concept of variety, resulting in path dependency, in spatial concentration and in specialization.
Another expected growth generator in EEG has to do with size: the larger the area, the larger the expected growth. However, this is bounded as it has the negative effect of wages increasing with magnitude (Frenken & Boschma 2007).
The concepts of EEG show that economic activity is geographically bounded (path dependence and routines) and subject to variety. Variety at a small level, which can be seen as related variety, results in economic growth on a small scale through economies of scope. Variety at the urban level, which is the same as unrelated variety, gives rise to economic growth through spillovers at a larger scale.

1.3 Urban Economics

Urban economics uses economic analysis to study urban areas. Urban economics studies the urban spatial structure and the location of households and firms (Quigley 2008). Unlike neoclassical economics, it also looks at relationships between individuals and organizations. It uses these relationships to explain what causes the formation and development of urban areas. Two important theories related to urban economics will be discussed, namely urbanization theory and agglomeration theory.

1.3.1 Urbanization Theory

Urbanization is, as the word already suggests, the transformation from rural areas to urban areas. Theories of urbanization have existed for a very long time and they are mixed with several other theories concerning urban growth. At first, industrialization was seen as the driver behind urbanization. Later modernization theory was related to urbanization, where technology and cultural diffusion became more important. The modernization theory was followed by dependency theory as modernization failed to account for developing countries. Urbanization in developing countries is a major spatial outcome of global capitalism and the countries' spatial organization according to dependency theory (Peng et al, 2010). It has a certain overlap with the path dependence and routines of the EEG. Regional economic growth through urbanization is expected to be dependent on the current state of urbanization in the region.
According to Friedmann (1986) world cities, which were the result of urbanization, exercise worldwide control over production and market expansion. World cities are major sites for the concentration and accumulation of international capital. They function as highly concentrated command points in the organization of the world economy and key locations for finance and specialized services, which have replaced manufacturing as the leading industry. They also function as innovative sites of production in these leading industries and as markets for the products and innovations of these industries (Sassen, 2001).

Urbanization theory maintains that the largest cities experience the highest growth, they are said to control the world. Economic growth is again expected to be unbalanced as was also put forward in earlier discussed theories. The regions where these world cities are located are very different from other regions and should also be treated differently to enhance economic growth.

1.3.2 Agglomeration Theory

As stated earlier, concentration of activities lowers transportation costs (Weber, 1929). The term used for this concentration of firms in a relative small area is agglomeration. The so-called clustering of firms results in linkages between firms where internal and external advantages are enjoyed. Marshall (1890) discussed three types of positive advantages for agglomeration. These are the lower costs for transportation of goods, people and ideas. All of these costs are lower when firms are located closer to each other. Firms can save shipping costs to suppliers and costumers, they can make use of a large supply of labor or even share labor and they can profit from intellectual spillovers (Marshall, 1890).

After Marshall, several studies have been done to validate Marshall's theory on agglomeration. Sometimes researchers tried to find out which of the three benefits is most important or they focused on one of the three benefits. Fujita et al. (1999) for example argued that the main driver behind agglomeration is the reducing of transportation costs for goods. Other researchers (Ellison et al., 2007; Audretsch & Feldman, 1995; Rosenthal & Strange, 2001) found evidence that agglomeration effects exist, but they saw no specific evidence that one of the three is most important. On the labor side it is found that in agglomerated areas workers can switch employers more easily, resulting in more productivity and reduction of wage differences (Krugman, 1991b and Diamond & Simon, 1990). The last type deals with transportation of ideas. Agglomerated areas experience intellectual spillovers; advantages arise when firms co-locate and workers learn skills from each other. Identifying this type of advantages is harder for researchers compared to goods and labor (Ellison et al., 2007). Most researchers use patents, but that is by far not the best way to measure intellectual spillovers as they occur between consumers and suppliers, which is better captured by input-output relationships (Porter, 1990).

Agglomeration, as stated above, provides benefits arising from locating near each other; the more economic activity in a region, the more benefits will arise. The other side is that there is more congestion and pollution and knowledge spillovers are geographically limited (Jaffe et al., 1993). This geographical limitation can explain the unequal distribution of growth across Europe (Romer, 1990). Caniels (2002) agrees with this and finds that knowledge spillovers are influenced by geographical distance and technological distance. In regional economic growth differences, spillovers are relatively more important than in economic growth differences on a larger scale. Regions are by definition smaller than countries and therefore interaction between regions is expected to be relatively higher. Regions are therefore expected to be spatially dependent of each other. Agglomerated areas thus experience knowledge spillovers. Combining this with new growth theory, which states that growth can also be endogenous, results in the knowledge spillover theory where knowledge created endogenously leads to more knowledge spillovers (Acs et al, 2008).

When linking agglomeration theory and regional economic growth two strands of literature need to be distinguished according to Beaudry and Schiffaurova (2009). On the one hand, it is argued that only as a result of specialization in a region knowledge spillovers can be enjoyed. Knowledge can only be transmitted in the same region and industry. Spillovers across different industries are not expected (Marshall, 1890; Romer, 1986). This type is also called localization or Marshall externalities. On the other hand, there are the earlier discussed Jacobs externalities, which hold that variety leads to economic growth. The more diverse or varied the region is, the larger the benefits (Jacobs, 1969).

Marshall externalities are expected to cause small, incremental innovations through the knowledge spillovers, while Jacobs externalities are expected to cause radical innovations through the spillovers. The smaller innovations tend to foster productivity growth of the firms itself, whereas the radical innovations provide employment growth, due to new combinations in products and/or technologies (Frenken et al, 2007). The next part elaborates further on variety and economic growth.

1.4 Variety and Economic Growth

Variety is earlier explained as a concept mainly used in mathematics to describe dispersions. In economics, variety can be used as product variety, but when focusing on regional economic performance one can have a broader look at variety. It is not per se the product variety, but rather the sector variety, which is important to examine. Of course there is an ideal level of variety that maximizes economic growth for every region, but unfortunately there is no widely accepted theory yet which can be used to come to this ideal level. Although theory linking economic growth and variety is missing so far some scientists have tried to examine the relationship. These attempts are put down further in this section where the two most interesting links are discussed in more detail.
The first one has to do with the agglomeration theory, endogenous growth theory, new growth theory and EEG. In all of these theories spillovers play an important role in economic growth. EEG even states that more variety at the urban level results in more spillovers, so automatically more variety has a positive effect on economic development (Jacobs, 1969; Glaeser et al., 1992; Van Oort, 2004).
A second way to link variety and economic growth comes from the portfolio theory. The portfolio theory is a concept used in business economics, where it shows that diversification helps reducing risks (Markowitz, 1959 and Montgomery, 1994). Diversification makes investments more stable and less vulnerable for shocks, which is of course very important for investors. The same theory can be applied on the relation between variety and economic growth. Variety protects regions for (big) unemployment growth when the economy is in bad weather (Frenken et al, 2007). This is in particular the case for unrelated variety, because in related variety an external shock can still have a big impact. A car manufacturer and a tire manufacturer, two firms that are related because they operate in the same sector, are both influenced by a large shock in oil prices, whereas a car manufacturer and a clothing shop, which are unrelated, tend to react in a different way to external shocks. Unrelated variety is thus supposed to protect regional economic development against external shocks (Attaran, 1986 & Haug, 2004). Unrelated variety can therefore protect against rises in unemployment due to external shocks. Related variety on the other side can enhance employment growth (Frenken et al., 2007).
Variety is also important in employment as it facilitates jobs for everyone, whether you are an engineer or a teacher. The chance of people finding their perfect job in their own region rises, as variety is larger. Following these theories, variety has a positive effect on employment growth and a negative effect on unemployment growth. For unemployment growth, unrelated variety is expected to have a larger effect than related variety.

1.5 Spatial Regimes

In line with the earlier discussed place-based development and smart specialization it is expected that different sets of regions (regimes) should be treated differently to enhance economic growth. This is due to the fact that European regions are not homogeneously distributed. The dataset gives a good opportunity to test for several regimes that are expected to explain this heterogeneous distribution. Four regimes are discussed; the size of cities within a region, European objective-1 regions, top university regions and capital regions. These four regimes and their expected influence will be discussed next. The output of the empirical research can provide more insight in the action and influence of the two concepts of place-based development and smart specialization, and it can show whether the tested regimes differ significantly from each other.

When regimes are significantly different, this means that different factors are important in causing economic growth. The output can also give an indication of the factors regional policy should focus on to generate economic growth in specific spatial characters of the regions. With specific spatial characters, the four regimes that will be discussed in the next part are meant.

1.5.1 Size of Cities within Regions

The European place-based development strategy as described in Barca et al (2012) argues that small and medium-sized regions with a polycentric structure and specialization in their best practices induce optimal economic performance as a result of the urban variety combined with specialization. The people-based or place-neutral strategy as described in the 2009 World Bank Report (World Bank, 2009) argues that the largest cities and metropolitan regions are the main forces behind economic growth. Recall from the urbanization theory that world cities are major sites for the concentration and accumulation of international capital. Large city regions are thus expected to experience higher economic growth compared to medium-sized and small regions.
In line with this place-based/place-neutral debate a regime is tested that has to do with the size of cities within regions. Regions containing large (at least 3 million inhabitants), medium-sized (between 1.5 and 3 million inhabitants) and small (lower than 1.5 million inhabitants) cities are disassembled and tested separate from each other. These groups are comparable to the OECD-distribution on a global scale (Boschma & Van Oort, 2012).

1.5.2 European Objective-1 Regions

Jacobs externalities as well as MAR spillovers contribute to economic growth according to the earlier discussed literature. Beaudry and Schiffaurova (2009) concluded in their meta-study on these two drivers that both specialization (Marshall) and diversity (Jacobs) play a positive role in economic growth. The reviewed literature shows a diverse picture of possible conditions and circumstances under which each driver could be at work. The choice of different economic performance measures, such as productivity growth and employment growth, and specialization and diversity indicators are reasons for different outcomes in different studies. Furthermore the levels of industrial and geographical aggregation play in important role. A regime that connects to the differences in industrial aggregation is the regime concerning objective-1 regions. Objective-1 regions are regions that received funds in the period 2000-2006 from the European Union through the European Regional Development Funds program. Regions lagging behind, having a GDP below 75% of the EU average, are supported through this program, which aims for cohesion in the European Union. Recent studies by Dogaru et al. (2011) and Marrocu et al (2012) concluded that objective-1 regions grow in productivity through specialization, while other regions grow in employment through diversity. In line with these recent studies it is expected that objective-1 regions experience higher productivity growth through specialization and other regions experience employment growth through diversity.

1.5.3 Top University Regions

The new growth theory says that regions with a larger stock of human capital experience faster growth through higher productivity. To be competitive in the globalizing knowledge economy nowadays, the local availability of knowledge and skills has become increasingly more important. Production is more and more heading towards value-added segments and knowledge intensive products and services (OECD, 2007).
Top universities are one of the possible actors that can attract these knowledge and skills into the region. The smart specialization strategy of the European Union addresses this with a guide to help improve the contribution of universities to regional development (Smart specialization platform, 2011). Regions containing top universities are thus expected to experience higher economic growth, especially through productivity growth.

1.5.4 Capital Regions

As said earlier centralization contributes to divergence according to Porter (2003) and Kim & Law (2012). According to Neoclassical growth theory, convergence is expected between European regions. But regional convergence at a European scale does not mean regions within countries are also converging. In fact, as said earlier, several studies concluded these differences on a European- and country scale (Cornett & Sorensen, 2010; Martin and Sunley 1998; Puga, 2002; Frenken & Hoekman 2006). With centralization policy, countries focus more on their central regions instead of remote regions. Centralization policy favors capital cities. These capital cities have become world cities due to this centralization and corresponding urbanization. Regions containing capital cities thus experience higher growth rates (Frenken & Hoekman, 2006). This centralization is expected to be one of the main forces behind the convergence/divergence issue on European and country scale. Capital regions are expected to show divergence with other regions in the same country, as they experience higher growth rates. When looking at capital regions versus non-capital regions in Europe it is expected that the aggregation of these two contribute negatively to convergence.

2. Method Section

No research has been done before on the impact of specialization, and related and unrelated variety on economic performance in regions on a European scale. This means that there is no relevant information available to investigate the influence of these variables without empirical research. Therefore empirical research is set up to test for these relationships. First, it is explained how economic growth is measured and which geographical scale is used. Second, the research model and corresponding conceptual model are discussed followed by the extensive description of all variables used in the empirical research. After this data overview the methodology is described.

2.1 Measuring Economic Growth

Economic growth is conventionally measured by the change in real GDP. Real GDP means gross domestic product adjusted for inflation. GDP or regional GDP is the sum of the value of all produced products in a specific country or region. The two most important components in creating regional GDP are the number of people employed in the region and thus the number of people contributing to the GDP in that region and their corresponding productivity. These two components eventually determine the regional GDP. In this thesis the two components are studied separate from each other. This way it is possible to assess for all the factors included in the research that possibly cause regional economic growth via which of the two main components they are of influence. Only taking into account regional GDP growth gives a more narrow view on what causes economic growth, whereas considering both employment growth and productivity growth separately can give more insight in the determinants behind regional economic growth.

Apart from employment growth and productivity growth, unemployment growth is also incorporated in this thesis. Unemployment is negatively related to economic growth. Unemployment growth does not necessarily result in negative economic growth, but it surely dampens the growth. Employment growth and productivity growth are the two components for output growth/economic growth, unemployment growth is an indicator of economic performance.

2.2 Geographical Scale

Statistics are used to study regional economic performance. A region is a geographical bounded area. For regional European statistics, Europe is divided into so-called NUTS areas (Nomenclature of Territorial Units for Statistics). The NUTS classification is a hierarchical system for dividing up the economic territory of the EU. There are three levels of subdivision; NUTS-1 is major socio-economic regions, NUTS-2 is basic regions for the application of regional policies and NUTS-3 is small regions for specific diagnoses. For example; the Netherlands consists of 4 NUTS-1 areas (North-, West-, East- and South Netherlands), 12 NUTS-2 areas (the provinces) and 40 NUTS-3 areas (COROP-regions). For this thesis the NUTS-2 level is used. An advantage of using the NUTS-2 level instead of NUTS-3 is that the NUTS-2 level is more suitable for researching the effect of unrelated variety as described in portfolio theory. At the same time, the EEG says that related variety works best at short geographical scales, while agglomeration theory suggests externalities work over a larger scale. As said earlier, Beaudry and Schiffaurova (2009) concluded in their meta-study that geographical aggregation was one of the reasons for the different outcomes of their reviewed literature. Because knowledge spillovers are geographically limited (Jaffe et al., 1993) and all levels have their pros and cons, for this thesis this medium is chosen.

The NUTS-2 regions of the following European countries are included in the research: Belgium, Czechoslovakia, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Slovakia, Spain, Sweden and the United Kingdom.

2.3 Research Model

The research model consists of six hypotheses, which were formulated taking into account the existing literature on the topic discussed in the first chapter and provides a conceptual framework to visualize the hypotheses. These hypotheses are listed in the next part. The first hypothesis deals with the expectation that regions are spatially dependent from each other. In previous studies, it is argued that interregional connections and spillovers lead to economic growth and the initial situation of regions has significant influence on performance. It is expected that high growth of a region is positively related to growth in neighbor regions.

The neoclassical growth theory states that convergence is the result of diminishing returns. Regions lagging behind in performance can catch up with more successful regions. This convergence is also supported by cumulative causation and new growth theory; economic growth is dependent on initial conditions and interregional interactions, where regions lagging behind can benefit from spillovers of neighbor-regions. Cumulative causation on the other hand supports divergence when economic performance is in an earlier development phase. This convergence issue is dealt with in the second hypothesis. The two components of output growth, namely employment growth and productivity growth, are expected to have negative relation with the initial state of the same components. Note that for unemployment convergence is not expected. Unemployment does not directly contribute to economic growth like employment and productivity; it is rather an indicator for economic performance.

The third hypothesis deals with the level of knowledge available in a region; the so-called human capital. New growth theory states that knowledge can be endogenous and more knowledge results in economic growth and knowledge spillovers. More education leads to more knowledge and therefore education is expected to be positively related to economic growth.

Agglomerated areas benefit from knowledge spillovers. In combination with new growth theory knowledge spillover theory arose, which states that knowledge created endogenously leads to more knowledge spillovers. As a result of specialization in a region knowledge spillovers can be enjoyed. These benefits are called localization externalities. Jacobs externalities in contrast to the localization externalities are benefits arising from a diverse or varied structure. Specialization is expected to foster productivity growth through small incremental innovations (hypothesis four), while variety is expected to provide employment growth through radical innovations (hypothesis five).

According to EEG spillovers are important, but also geographically bounded and subject to variety. Related variety results in economic growth on a small scale through economies of scope. Unrelated variety, gives rise to economic growth through spillovers at a larger scale. Variety has, through portfolio theory, a positive effect on employment growth and a negative effect on unemployment growth. For unemployment growth, unrelated variety is expected to be of greater influence than related variety. Hypothesis five is therefore split in two components of related variety and unrelated variety.

New economic geography says that economic performance is geographically unbalanced and concentrated in few densely populated areas. Urbanization theory supports this heterogeneity by concluding that the largest cities experience the highest growth. Hypothesis six incorporates these expectations. Policies for enhancing growth need to be region-based. The different regimes discussed in chapter 1.5 contribute to this expectation, as all the regimes are expected to have significant influence on the model.

In hypothesis two and four regime-specific expectations are included. In hypothesis two the capital region regime is expected to have a positive effect on convergence, because centralization is said to result in divergence between capital regions and non-capital regions. Objective-1 regions in hypothesis four are expected to experience even higher productivity growth through specialization than other regions due to the support they get.

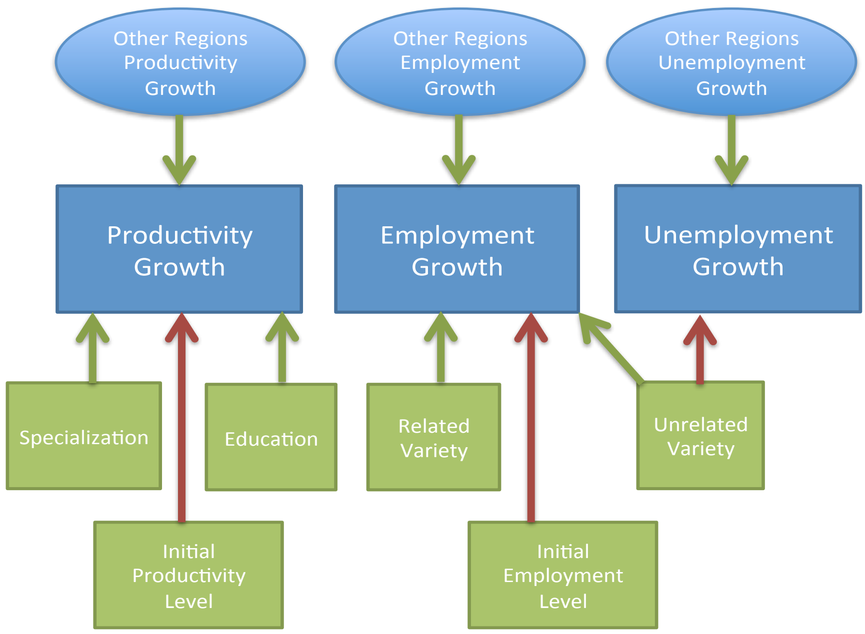
2.3.1 Hypotheses

1. European regions are geographically dependent upon each other for their economic growth; growth of neighbor regions is positively related to growth in the own region.
2. There is convergence in employment and productivity between European regions, in particular when distinguishing between capital regions and non-capital regions.
3. The education level is positively related to economic growth.
4. A specialized sector structure has a positive effect on economic growth via productivity growth
    a. Specialization is positively related to productivity growth, even more in objective-1 regions
5. A varied sector structure has a positive effect on economic growth via employment growth and negative unemployment growth.
    a. Related variety is positively related to employment growth.
    b. Unrelated variety is positively related to employment growth and has a stronger negative effect on unemployment growth.
6. Economic activity is particularly heterogeneous and therefore regions should be treated differently according to their profile to enhance economic growth.

2.4 Conceptual Model

In the conceptual model (figure 1) most of the hypotheses are put together and visualized. The circles at the top show hypothesis one, concerning the spatial dependence of the regional performance. They all have positive influence on the dependent variables indicated by the green arrows. The two longer red arrows from the green boxes with the initial performance levels to the dependent variables visualize the second hypothesis, which deals with the convergence issue. The red colour of the arrow implies a negative relation. For unemployment no convergence is expected as explained earlier. Hypotheses three, four and five are represented by the other green boxes, with their adherent arrows showing the type of influence.
Off all hypotheses only the sixth hypothesis was not put into the model, as well as the added regime conditions in the convergence hypothesis and the specialization hypothesis. To test for these regime-related expectations an extension of the model is made in which the dataset is split to see whether there are different outcomes for different groups of regions. Consider the conceptual model below to be duplicated for the different groups of regions; each group is put in the model separately. The expectation is that the models differ significantly from each other, which means that the groups are not similar in the influence of the variables on the dependents used in the model.

Figure 1: Conceptual model



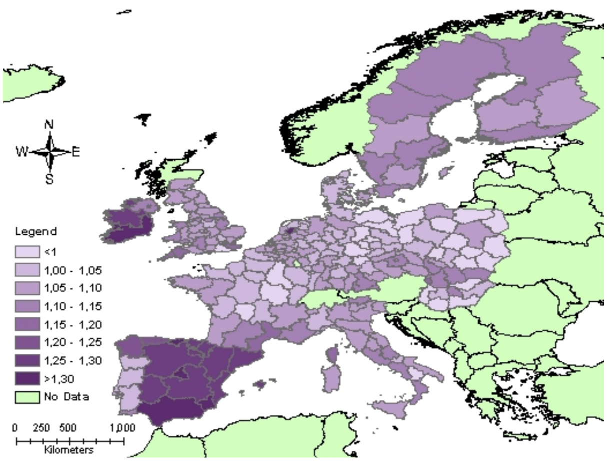
2.5 Data Overview

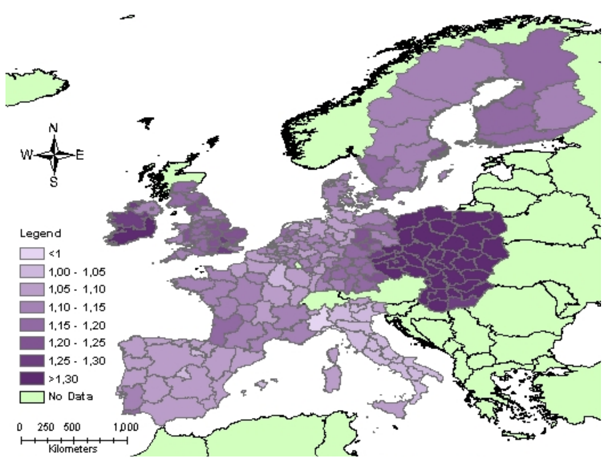
In this thesis, the focus is on regional structure and spatial dependence of the regions. Variables that say something about the structure of the region are therefore the most important independent variables. By using spatial analysis, the spatial dependence is captured. Eventually, all the separate models are used to explain regional economic growth.

2.5.1 Dependents

As stated earlier economic growth is conventionally measured as growth in GDP, which is a similar measure as the growth of total output in a region. Of course output growth is the starting point for the dependent variables, but to provide a more detailed overview, employment growth, productivity growth and unemployment growth are the main dependents, which are examined separately in the empirical research.
The data for employment and productivity are taken from the Cambridge Econometrics dataset. In the dataset each region has a figure for their employment and their productivity in the years 2000 and 2010. The easiest way to calculate growth is by taking the percentage change from 2000 to 2010, but in this research growth is calculated in a different way. The rationale behind this is that growth can also be negative and because the data has to be transformed to natural log the percentage change cannot be used, because taking the natural log of a negative figure is not possible. By using the "new" divided by "old" method this problem is tackled, because it gives a similar view on growth as the percentage change. Yet the outcomes are now around "1" instead of "0", where figures below 1 mean negative growth and above 1 mean positive growth.
Figures 2a and 2b give an insight in the growth levels of each region. It is clear that both maps show country patterns. In employment growth Spanish and Irish regions are performing relatively well, while in productivity growth the Irish and Polish regions show the highest growth rates. Based on these two maps there is clearly spatial dependence between the regions, because regions that are close to high growth regions tend to show high growth themselves also.

Figure 2a: Employment Growth (2000-2010) Figure 2b: Productivity Growth (2000-2010)





For calculating the output, the employment level and corresponding productivity level of each region in the years 2000 and 2010 are multiplied, which gives the 2000 and 2010 output level. Output growth is than calculated in the same way as employment and productivity growth using the 2000 and 2010 figures. Figure 2c shows the output growth of all the European regions in the dataset used. Of course Ireland is performing quite well in output growth, because output is a result of the combination of productivity and employment and figure 2a and 2b show high growth in Ireland for both of these components. Most Polish regions show high growth rates and Spanish regions are also doing very well.

Based on figures 2a, 2b and 2c it can be concluded that the relatively high output growth in Poland resulted from the high productivity growth and the output growth in Spain was a result of the high employment growth.
Figure 2d gives insight in the distribution of unemployment growth in the European regions. The data for unemployment are obtained from Eurostat. Again the same method is used for calculating the growth, but now growth is measured as the 2010 level divided by the 2003 level. This is an even shorter period of seven years compared to the time span of ten years used in the other dependents. Unemployment growth shows country patterns as well. Denmark, Hungary, the United Kingdom and Ireland show high unemployment growth, just like some regions in Portugal, Spain and the north of Italy. Low unemployment growth is particularly visible in Polish and German regions.

Figure 2c: Output Growth (2000-2010)

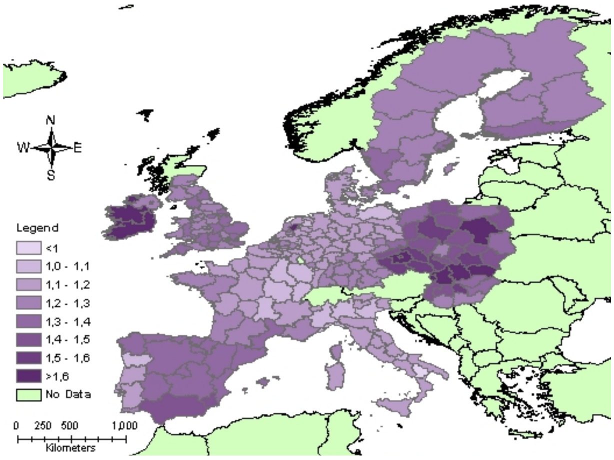
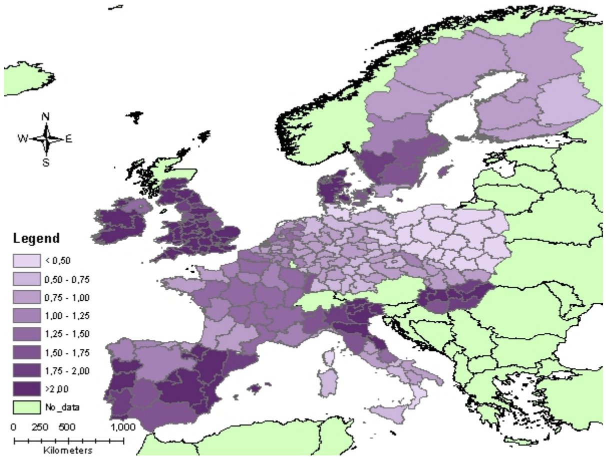
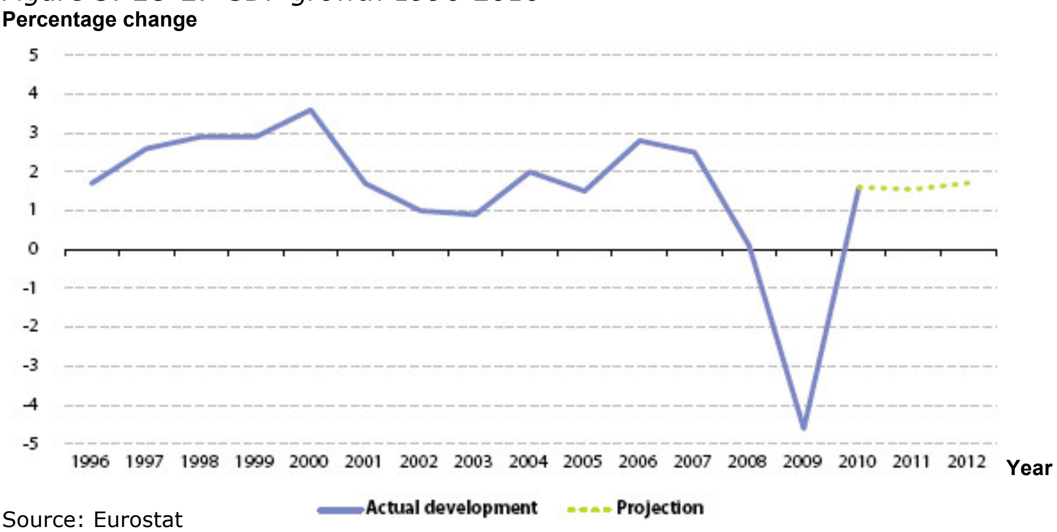


Figure 2d: Unemployment Growth (2003-2010)



One of the main remarks for the calculations of growth is that only one period is taken to come to the growth figure of each region and also that this period is only ten years, which is quite a short period that could have been influenced by cyclical changes in economic performance. The financial crisis for example might have a big influence on the model. Figure 3 shows how GDP growth evolved on average from 1996 to 2010 in the total EU-27 area. Two different periods of ten years, 1996-2006 and 2000-2010, have a very different average growth figure due to the financial crisis. The highest growth figure was 3,6% in 2000, while the lowest growth was measured in 2009 with -4,6% growth. It is unlikely that these fluctuations are explained by the independents used in this thesis, but it definitely matters which period is chosen as the average growth may be influenced substantially due to the economic crisis.

Figure 3: EU-27 GDP growth 1996-2010

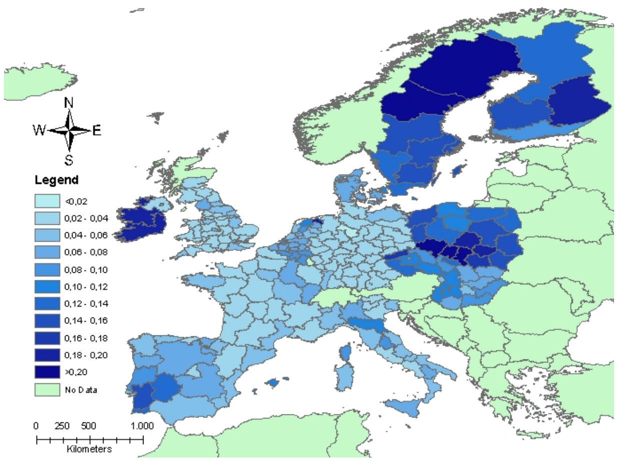


Source: Eurostat

2.5.2 Independents

The three main independent variables for testing the hypotheses in this thesis are all related to the sector structure. The first variable captures the degree of specialization in a region. The degree of specialization is calculated as the Theil index over the location quotients of 14 sectors, as distinguished in the Cambridge Econometrics data (see Dogaru et al. (2011) for more detailed information on this calculation). The dataset is collected by the Netherlands Environmental Assessment Agency (PBL) and is based on regional production and trade data for all the 205 NUTS-2 regions in this thesis. It is thus a combination of the Theil coefficient and the location quotients of the region. The Theil coefficient measures deviations from the European average distribution of production specialization in all sectors and the different location quotients measure the relative specialization of a region in a certain sector. A high score represents a large degree of sectoral specialization in a region, and a low score represents sectoral diversity of a region (Dogaru et al, 2011). Figure 4 maps the specialization level of all the 205 regions used in the dataset. Specialization is relatively high in Sweden, Finland, Ireland, the South of Portugal and some regions in Eastern Europe. Regions in Germany, England and France show low levels of specialization. Regions in these countries are thus relatively more diversified, because diversity is the counterpart of specialization.

Figure 4: Specialization



The two other important independents for this thesis, related and unrelated variety, have to do with the diversity of economic activity in the different regions. Diverse regions have high variety and as discussed earlier this variety can be split into related and unrelated variety. In this empirical research both types of variety are examined using a dataset of firm activity in all European regions. This dataset is called Amadeus and is made by Bureau van Dijk. The dataset is a sample collection of the financial statements of all individual firms in most of the European countries made by Bureau van Dijk. Data is available in four so-called Nace levels. Nace is a European industry standard classification system consisting of a 6-digit code. The first four digits of the code (the first four levels of the classification system) are the same in all European countries. The fifth and the sixth may differ from country to country. The first digit distinguishes in larger economic sectors, the second in subsections or divisions, the third in groups and the fourth in classes. The more digits, the more detailed is the classification. The Amadeus dataset shows for every region the economic activity specified on the Nace-1 through the Nace-4 level. It thus gives insight in the spread of economic activity in several classifications.

Two types of data are available, namely the number of firms and the sales volume in each region. Both datasets can give a different indication of variety. The number of firms does not distinguish between very small and very large firms, while the turnover gives a better view on the regional structure. A region with one very large firm in a specific sector and several small firms in other sectors is concluded to have higher variety than actually true (Witte, 2011). Therefore, this thesis uses the weighted dataset with turnover quantities instead of the dataset that looks at the number of firms.

By using this dataset variety can be calculated on every Nace-level using an "entropy" formula (Frenken et al., 2007). The entropy formula is a formula used to calculate variety at every possible scale. One can calculate for example the variety in an entire country, in a region or in a city. Apart from these geographical scales, calculating variety within a specific sector in a specific region is also possible.

Entropy measures the spread of activity for different cases within a specific population. It measures to what extent the population differs from being perfectly homogeneous. When activity is spread perfectly over every sector, within a specific region, entropy is maximized and when activity is only in one sector entropy is zero, thus the higher the entropy the more variety. The maximum value depends on the number of sectors within the region, where an additional sector means a higher possible entropy score. Equation 1 shows how entropy E for region j is calculated (adapted from Frenken et al., 2007: 689, equation 1):

Equation 1: Entropy formula

$$E_j = \sum_{k=1}^{n} \frac{firms_{jk}}{firms_j} \log_2\left(\frac{1}{firms_{jk} / firms_j}\right)$$

One of the main advantages of using entropy in measuring variety is that it can be compared over different regions and that different calculations are possible. It is for example allowed to sum entropy in different sectors in every region and subsequently compare these numbers for every region.

As said earlier unrelated variety is variety at a larger scale, while related variety is variety at a smaller scale. Unrelated variety can thus best be measured by looking at variety at the largest possible scale. The largest scale at which variety can be calculated is by calculating entropy at Nace-1 level within the regions. Unrelated variety of region j can be calculated using equation 1. Unrelated variety (UV) is than calculated as entropy of economic activity in 21 broad sectors (k) within the region (j) as shown in equation 2.

Equation 2: Unrelated variety

$$UV = \sum_{k=1}^{n} \frac{firms_{jk}}{firms_j} \log_2\left(\frac{1}{firms_{jk} / firms_j}\right)$$

Entropy at Nace-2 level can be calculated with the same formula, but than 86 subsectors are used instead of 21 broad sectors. The most important reason why the Nace-1 level is used for calculating the unrelated variety instead of the Nace-2 level (compare with Frenken et al, 2007) is that variety at Nace-2 level gets closer to related variety. Extraction of coal and extraction of iron ore for example are separated into two different Nace-2 subsectors, while they are put together at Nace-1 level in the sector mining and quarrying. Distribution of activity between coal and iron extraction is not unrelated variety, but rather related variety. Distribution of activity at the Nace-1 level is most suited for calculating unrelated variety, as it looks at distribution between 21 broad sectors like industry, mining and quarrying, construction, education etcetera.

Related variety is more complex to calculate, because related variety is variety at a smaller scale, which means that the variety of numerous classes needs to be calculated for every region. Frenken et al. (2007) calculates the related variety of the regions as the weighted sum of variety at the Nace-5 level within each Nace-2 division. It is argued that related variety is best measured as the entropy at the 5-digit level within each 2-digit class. See for further explanation Frenken et al. (2007).

Because the 5-digit level is not comparable over European regions, it is better to use the 4-digit level instead. The Nace-4 level is also the smallest available classification in the Amadeus dataset. While Frenken et al. (2007) use weighting at the Nace-2 level, this thesis makes use of unweighted related variety. With weighting at the Nace-2 level the 4-digit entropy in a large Nace-2 division has more influence on the figure for total related variety in a region than 4-digit entropy in a small Nace-2 division. A five million euro change of turnover from one Nace-4 class to another Nace-4 class within a larger Nace-2 division therefore has a bigger influence than a five million euro change of turnover at Nace-4 level in a smaller Nace-2 division. In other words; changes in the distribution of activity at the 4-digit level within each Nace-2 division have different influence on the total related variety figure of the region depending on the size of the Nace-2 division in which this change occurs.

Take for example a simple region with activity in only two divisions. Initially, the first division has €10 turnover in the entire division, all in one 4-digit class while the second division has €90 turnover in one 4-digit class. Entropy in both divisions is zero and thus total related variety in the region is zero using both methods.
Now, only in the first division €5 of the turnover moves to a different 4-digit class, resulting in two 4-digit classes with both €5 turnover. This clearly means more related variety in the first division, actually at first there was no related variety and now there is. Total related variety is now 0,10 using the weighted method and 0,50 using the unweighted method.
Next, in the second division €5 of the turnover moves to a different 4-digit class, resulting in one 4-digit class €85 turnover and another with €5 turnover within this second division. Total related variety is now 0,47 using the weighted method and 0,71 using the unweighted method. In the weighted method, related variety more than quadrupled while in the unweighted method related variety increased by 42%. It is clear that a €5 change in turnover in a larger division results in substantially larger related variety using the weighted method. While division one has a perfect distribution of 50% in each of the two 4-digit classes division two has an unequal distribution of €5 in one class and €85 in the other. The entropy in division one is 1,00 and in division two 0,42. It is clear that the change in class two shouldn't be valued higher for the total related variety than the change in turnover in class two.
In fact, it should be valued less as the shift in division one is much more radical for the division than the shift in division two. The unweighted method does the best job here as related variety changed by 0,50 after the shift in division one and 0,21 after the shift in division two. In the weighted method total related variety is composed by 0,10 as a result of the shift in class one and 0,37 as a result of the shift in class two. This example shows clearly that the unweighted method is a better way to measure related variety. Changes with the same pecuniary magnitude in smaller divisions are more visible with the unweighted method, while they are almost invisible in the weighted method. The weighted method actually focuses mostly on the related variety in the largest divisions. The weighted method also does not account for divisions without any activity, because they simply get no weight, while the unweighted method does account for these divisions. No activity in a division is also important information and accounting for this makes it fairer to compare related variety figures across regions, as related variety is likely to be spread differently across divisions in every region.
The weighted method on the other hand, accounts better for related variety in divisions with relatively more activity. In larger divisions more spillovers can take place, as there are simply more firms to transfer knowledge and ideas to. There is thus something to be said for both types of calculations.
Whether the weighted or unweighted method reflects related variety best cannot be said with certainty. Actually the two methods show quite similar results as indicated by the moderate agreement between the results of the two methods (correlation is 0,507).
In this thesis though, related variety is calculated using the unweighted equivalent. Now recall from the earlier discussion regarding the measuring of unrelated variety that unrelated variety is measured as variety at the Nace-1 level because variety between these 21 broad sectors is considered to be unrelated while variety at Nace-2 level is considered to be more related. Variety within a Nace-1 division is considered to be related variety so for the most accurate measurement of related variety, related variety should be measured as Nace-4 variety within the broad Nace-1 sectors instead of within the Nace-2 divisions.
Related variety is calculated as the unweighted sum of entropy at the Nace-4 level within each Nace-1 division. First, for every Nace-1 sector (j) in the region the entropy at the Nace-4 level (k) is calculated. This gives a sum of numerous entropy scores (L) at 4-digit level within the 1-digit level for every region. Than, the sum of all these entropy figures in a specific region makes the related variety (RV) of that region. Equation 4 shows how this is finally calculated.

Equation 3: Entropy at Nace-4 level within a Nace-1 division

$$L = \sum_{k=1}^{n} \frac{firms_{jk}}{firms_j} \log_2\left(\frac{1}{firms_{jk} / firms_j}\right)$$

Equation 4: Related variety

$$RV = \sum_{L=1}^{n} L$$

Both related variety and unrelated variety can be compared from region to region. Yet, comparing unrelated variety with related variety is not possible, due to the different types of calculation.

Of course the two types of variety are related to each other as they both measure variety, but they do actually measure two different types of variety and correlation between the unrelated and related variety is rather low (0,237), so it is possible to put the two variables together in a model although they probably do take away part of each others' explanatory value.

Figure 5a and 5b show the variety levels in the European regions. The two figures make clear that related variety and unrelated variety are indeed different from each other. Related variety is high in the Belgian regions and in some regions containing large cities like the region of Paris, Barcelona, Madrid and Warsaw. Apparently, the size of the region matters in related variety. Densely populated and/or large areas show higher related variety figures. The variable population density controls for this issue and correlation between population density and related variety is only 0,306 so it can be concluded that this finding is not alarming.

Figure 5a: Related Variety

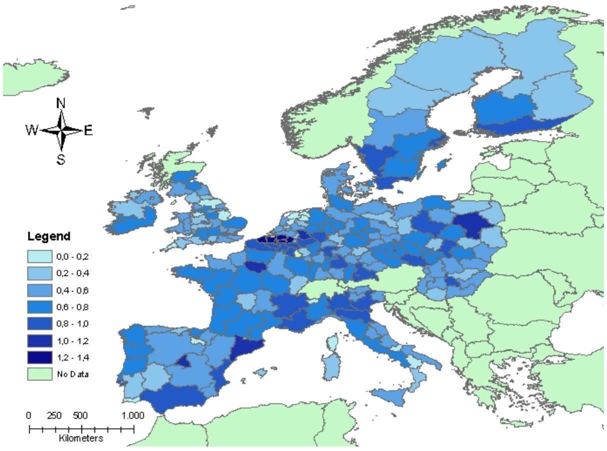
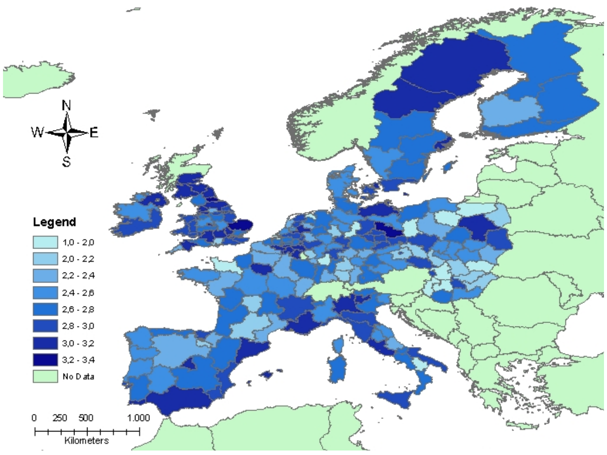


Figure 5b: Unrelated Variety



Unrelated variety is also higher in these "large city" regions, but also in some regions in Great Britain, the north of Sweden and the south of Spain for example. In both related and unrelated variety there is no real country pattern visible. High and low levels of variety seem to be spread over Europe quite equally.

2.5.3 Controls

Although there is a focus on specialization and variety in this thesis, there are several other factors that might play a significant role in unemployment, employment and productivity growth according to the literature. Gardiner et al. (2004) refer to several factors in its "pyramid model" of regional competitiveness and Van Oort (2004), Combes et al. (2008) and Brulhart & Mathys (2008) give extensive overviews as well.

When measuring growth, the most important control variable is likely to be the variable measuring the initial state of the dependent variable. Employment growth from 2000 to 2010 is likely to be influenced by the employment level at the year 2000. Growth is expected to be more difficult to achieve when a region already has relatively high employment. The same might hold for productivity and unemployment.

Other control variables used in this thesis are private and public research and development (R&D), openness, market potential, education, population density and the wage level. They are all calculated for the year 2000 just like the initial levels of the independent variables (except for unemployment which is calculated for 2003). Data collection on NUTS-2 regional level is described in more detail in PBL (2011) and Dogaru et al. (2011). Figures 6a and 6b show the levels of employment and productivity in the year 2000 across Europe. Figure 6c shows the unemployment level in the year 2003. Here the same data is used as for the calculation of the corresponding growth figures.

Figure 6a: Employment 2000 (millions)

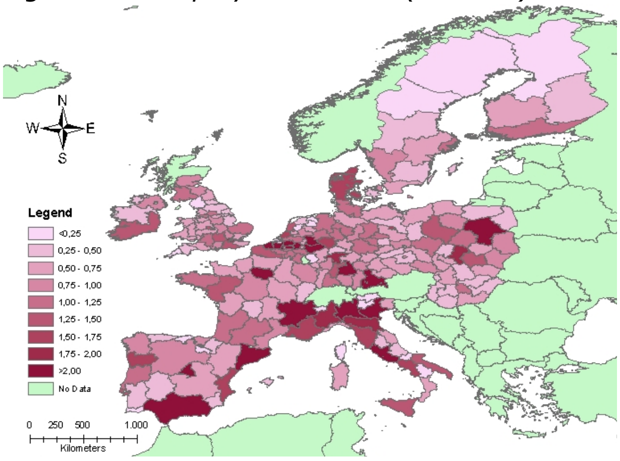
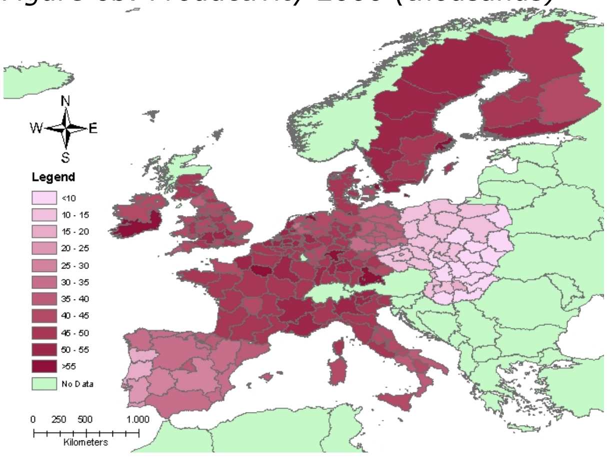


Figure 6b: Productivity 2000 (thousands)



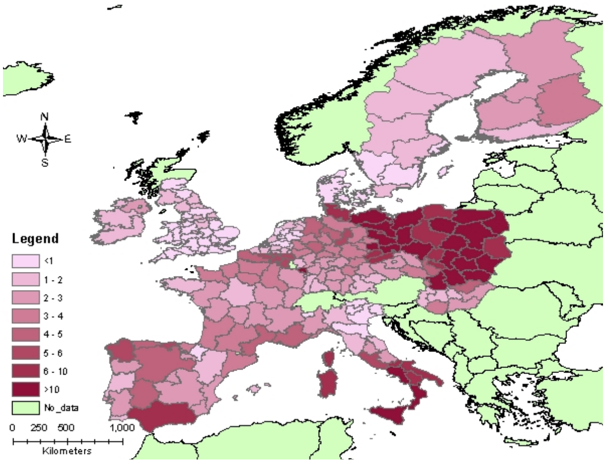
In figure 6a the employment level in regions in 2000 is mapped. Larger regions, capital regions and "large city" regions are showing high employment levels. This is not surprising as absolute figures are calculated that are definitely influenced by the size of the region.

Figure 6b shows productivity in 2000 in thousands of Euro's. It can be concluded that there is a pattern visible where the core economic area of Europe has high productivity and regions in Eastern European and Spain have lower productivity in 2000.

When comparing figure 6b with figure 2b it can already be concluded that expectations of convergence for productivity are true, because there are clearly contrary patterns visible. Initial levels are expected to have a negative influence on the corresponding growth levels. The figures in annex 1 provide an even better insight into how this relationship functions. Dependent variables are related with their corresponding independent variable and they make clear that for productivity convergence takes place. Whether convergence also takes place in employment cannot be concluded with certainty. In annex 1 the regions in different countries are indicated with colors, which shows again the country patterns that were already visible in figures 2a to 2d.

Figure 6c shows the unemployment levels in 2003 in % of the total workforce per region. As this is a percentage, the size of the region must have less influence on the scores, but again a pattern is visible. Regions in Poland and the south of Italy and some regions in Spain show relatively high unemployment scores.

Figure 6c: Unemployment 2003 (%)



Private and public R&D are calculated as the percentage of total GDP of the region spend on the private and public R&D respectively using Eurostat statistics. They are expected to have a positive influence on economic growth (Moreno et al., 2006). Figure 7a and 7b show the distribution of R&D spending in the 205 European regions. Notable is that regions in Sweden, Finland and the south of France are doing very well in both private and public R&D. Italy shows low expenditure in private R&D, but higher expenditure in public R&D. Private R&D is more visible in regions with larger multinational enterprises, while public RD is more attached to regions with technological universities and regions where universities and firms alliance (Dogaru et al., 2011).

Figure 7a: Private R&D

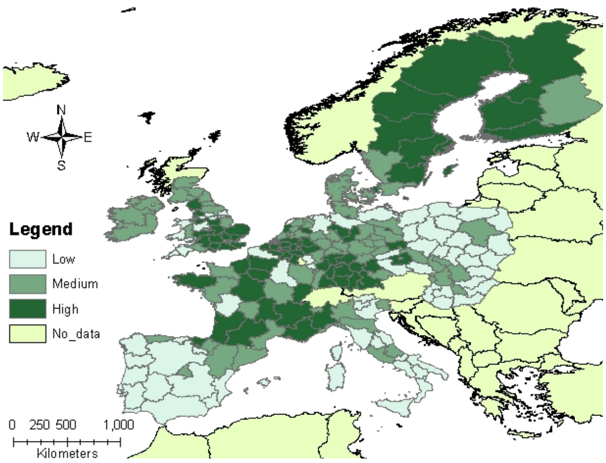
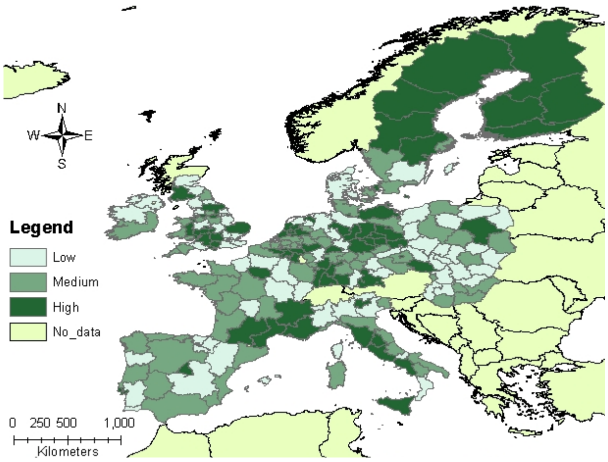


Figure 7b: Public R&D



Openness of the economy is deducted from PBL (2010). It is calculated as the total sum of import plus export as percentage of GDP. Market Potential is calculated using a gravity model on population and employment. Figure 7c and 7d demonstrate the openness and market potential of the dataset used. The openness clearly shows country patterns, as openness is relatively high in all regions in Portugal, Ireland, Slovakia, Hungary, Czechoslovakia, Sweden, Finland and the Netherlands. Openness can have a positive effect, as it attracts economic activity. However, it can also have a negative effect, as economic activity may spill over to other regions easier in more open regions.

Market potential also shows an interesting picture; here high numbers are centered and the further away from the center of the dataset the lower are the scores. It is expected that a higher market potential is positively related to growth rates due to larger market and customers' opportunities, potentially higher profits and more incentives for innovation and renewal (Dogaru et al., 2011). Both openness and market potential illustrate that the dataset is not random, but regions are connected because geographical patterns are visible.

Figure 7c: Openness

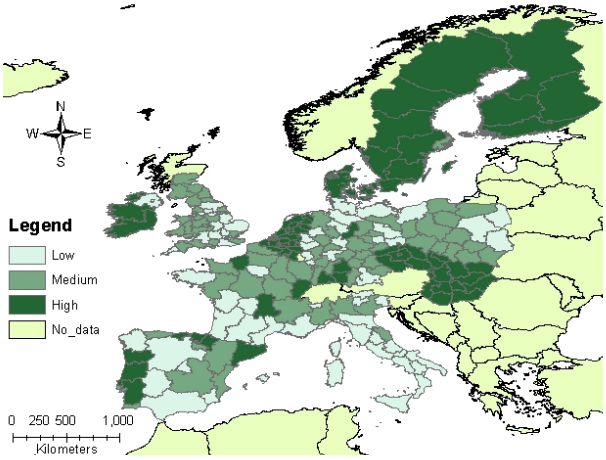
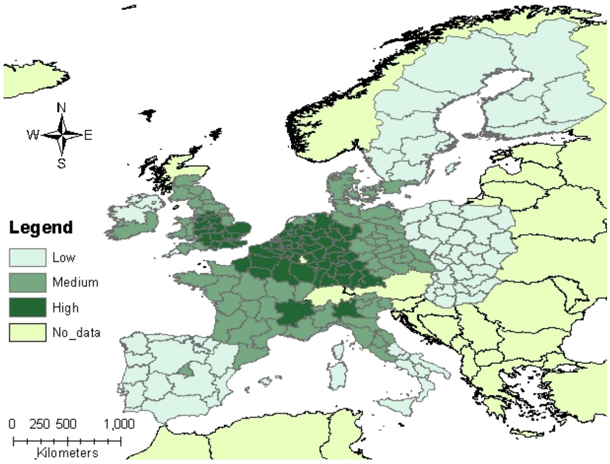


Figure 7d: Market Potential



Education is calculated as the share of high education with respect to total education. A region scoring high in education thus has a larger share of high education in total education. Figure 7e shows these scores where it can be concluded that regions in Eastern Europe, Italy and Portugal have smaller shares of high education. Like in openness and market potential, education also shows specific country differences. Education has an overlap with the wage level as visible in figure 7f. Here regions in Portugal and Eastern Europe also show low scores. Italian regions score medium on wage level though and some regions in France score low on education and high on wage level. Both education and the wage level are expected to have a positive influence on growth.

The last control variable is shown in figure 7g. This variable measures the population density in the regions. It controls for densely populated areas and is calculated as the total number of people living in a region divided by the size of the region. Regions with metropolitan areas like Paris, Rome, Madrid and London for example, show high population density while remote regions like the south of Spain and the north of Scandinavia show low population density. This is no surprise of course, but again it shows that the dataset is nested.

Figure 7e: Education

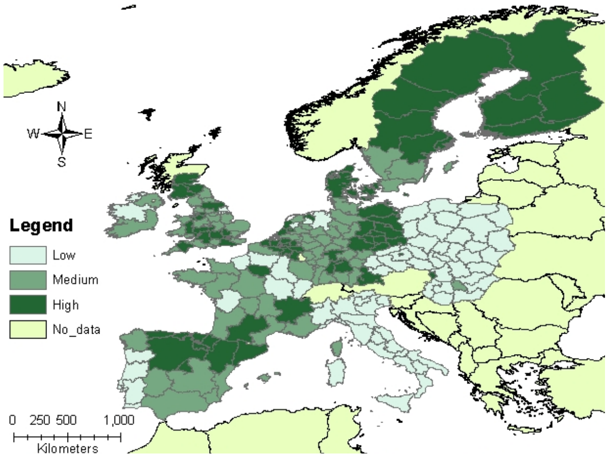


Figure 7f: Wage level

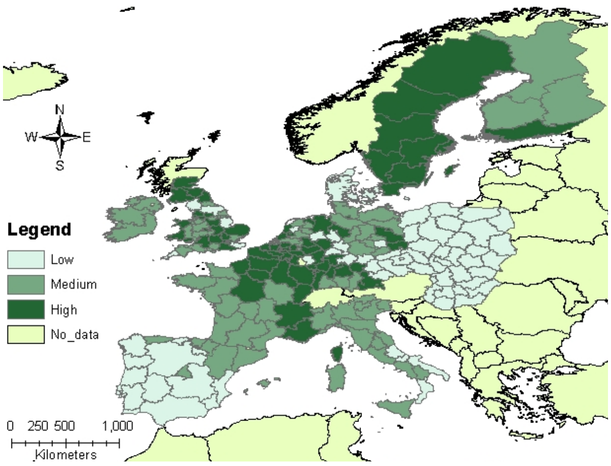
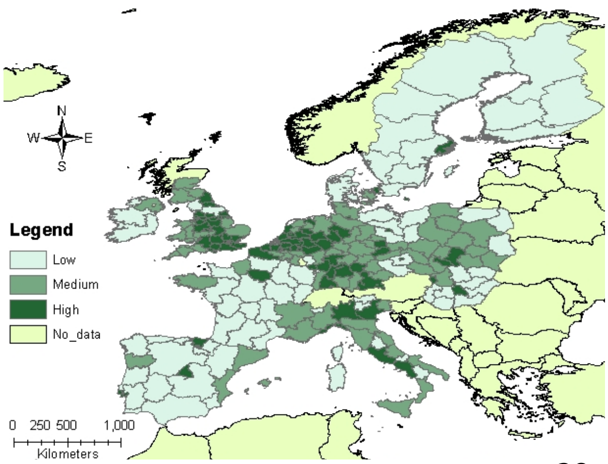


Figure 7g: Population density



2.6 Methodology

The dataset contains 205 connected regions in Europe. As a result of the nested structure, which has been shown several times , ordinary regression and correlation are potentially deceiving in the presence of spatial autocorrelation. The 205 NUTS-2 regions are all different cases and although they differ in size, every region has the same influence on the model. Unfortunately there is no easy, statistical or theoretical way to reconcile all the regions for their size. The methodology of performing spatial analysis as described later in some way helps reducing this problem. For larger regions the neighbours are further away and thus in these larger regions the own region is influenced less by other regions in the spatial analysis. Although spatial autocorrelation is very likely to be present, a simple OLS-model is conducted first. It is expected that due to the presence of spatial autocorrelation the results are biased and falsely significant (Anselin, 1992). Equation 5 shows the formula for this first test.

Equation 5: OLS-model

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots\ldots + \beta_{11} X_{11} + e$$

Y is the dependent variable, the X's are all the independents including the control variables with their corresponding influence indicated in the beta and alpha is the value of Y when all betas are zero. After this first model a spatial model is conducted which is likely to be a better fit. A spatial weight matrix is used to capture the spatial relations. At last, different regimes are tested in search for more explanatory power.

2.6.1 Spatial Weight Matrix

The question here is how to measure the spatial relationships. In this thesis the spatial analysis is performed using the geographical distance between the regions. Although relationships between regions can also be measured in travel distance or by networks of trade and knowledge, geographical distance is chosen as the spatial weight measure. The reason is that it is relatively simple to calculate and it gives a fair first view on the spatial relation between regions. The spatial model is an expansion of the OLS-model where a new variable is introduced namely W_Y. The new formula is shown in equation 6.

Equation 6: Spatial model

$$Y = \alpha + WY + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots\ldots + \beta_{11} X_{11} + e$$

The extra variable W_Y indicates the spatial dependence. It is the influence of other regions' Y to the own regions' Y. It is calculated as the sum of all the other regions' Y (RY) multiplied by the inverse distance (d) of that other region to the own region. Equation 7 shows the formula for the derivation of this spatial dependence.

Equation 7: Spatial dependence of the dependent

$$WY = \sum_{R=1}^{n} (RY * dR^{-1}) - Y$$

Every region influences the own region, and this influence is eventually added to the sum of the total influence of all regions to the own region. But of course as distance grows the influence gets smaller. Therefore the inverse distance is used. For example a region with a distance of 25 to the own region is multiplied by 25^{-1} = 0,04 and thus has more influence than a region with a distance of 100 to the own region (100^{-1} = 0,01). The further away the region is from the own region, the less influence it has on the own region. Distances are measured from the geographically weighted middle of the own region to the geographically weighted middle of the other region. Subsequently, equation 8 shows the regression formula for the distance weight matrix with power 2. Here, the influence of other regions reduces quadratic with their distance to the own region as shown in equation 9.

Equations 8 and 9: Quadratic spatial dependence model

$$Y = \alpha + W^2Y + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_{11} X_{11} + e$$

$$W^2Y = \sum_{R=1}^{n} (RY * dR^{-2}) - Y$$

This last model is expected to fit even better than the model shown by equation 6 because theories like EEG and agglomeration theory argue that spillovers are geographically bounded. The influence should therefore be reduced exponential rather than linear in the model. Take for example a region with a distance 10 to the own region and a region with a distance 20 to the own region. In the weight matrix with power 1 influence of the closer region is two times higher, while it is also two times closer to the own region. In the weight matrix with power 2 influence of the closer region is four times higher while it is still two times closer to the own region. The model with power 2 thus seems to be a better fit, as it accounts for the diminishing influence.

With this spatial analysis the spatial autocorrelation is reduced. This is important to improve the model fit, but it is also suitable to capture spillover effects, which are important in this thesis. When W^2Y turns out to be significant, than it can be concluded that there is something going on spatially, where a positive W^2Y indicates a positive influence of the other regions. Due to the inclusion of W^2Y it is expected that the other dependents become less significant. Part of the explanatory power is taken over by the spatial dependence. This is not a problem, because the spatial model is expected to be a better fit and it thus gives a better view on the real influence of the dependents.

2.6.2 Regimes

After the spatial analyses the data is also used for the effect of the different regimes discussed in chapter 1.5. In each regime, the dataset is split in two groups (or three groups for the size of cities regime) on the basis of the specific characters discussed in the literature review. The groups are tested separately using the same model. The resulting outcomes can be compared and a statistic test, named the Chow-Wald test, will show whether the outcomes differ significantly from each other. In all the analyses the spatial weight matrix with power two is used to capture the spatial autocorrelation.

The four regimes that are tested are shown in figure 8a to 8d. Figure 8a shows the map of the first regime tested in this thesis; it is the regime that looks at the size of cities within the regions. Regions containing small cities are separated from regions containing medium sized cities and large sized cities. This first regime thus divides the data into three groups rather than the other regimes where the data is divided in two groups. Figure 8b shows the regime where objective-1 regions are separated from the rest of the regions. The blue regions are the objective-1 regions.

In figure 8c, top university regions are marked blue. These regions contain at least one university in the Shanghai top-75 ranking of European universities. There are in total 47 regions with at least one top-75 university, so some of these regions might have two or even more top universities.

Figure 8a: Size of Cities in Regions

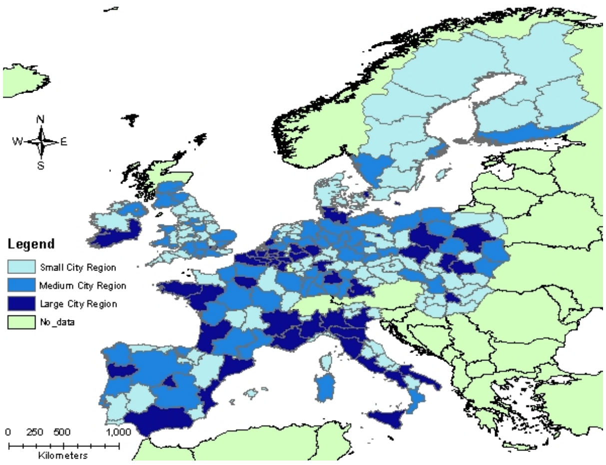
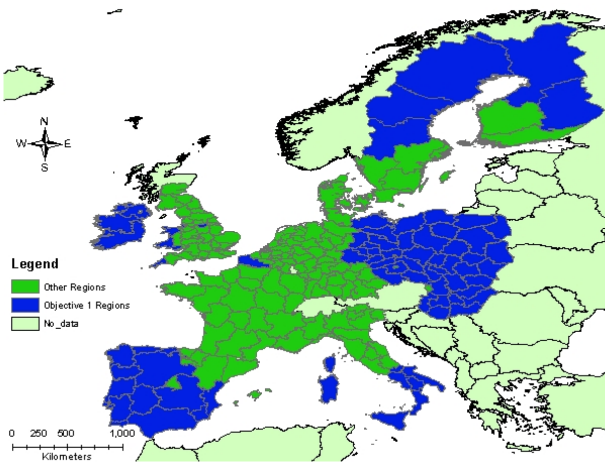


Figure 8b: Objective-1 Regime



The last regime has to do with regions containing a country capital. Due to the fact that the dataset contains 16 countries there are also 16 capitals. Almost all capitals are within one region but only one capital, London, is divided over two regions. This regime thus contains 17 regions marked as capital region. In figure 8d these regions are highlighted in blue.

Figure 8c: Top University Regime

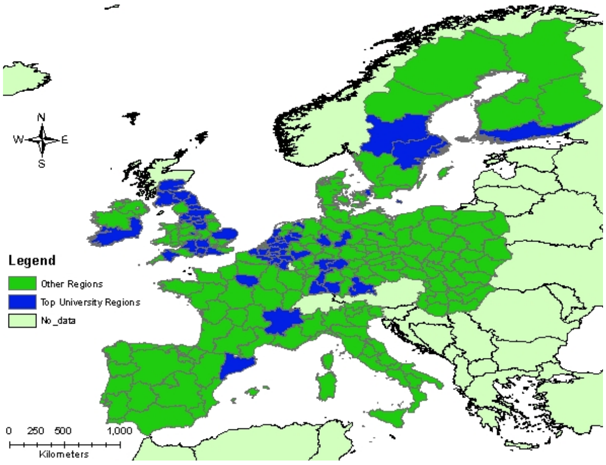
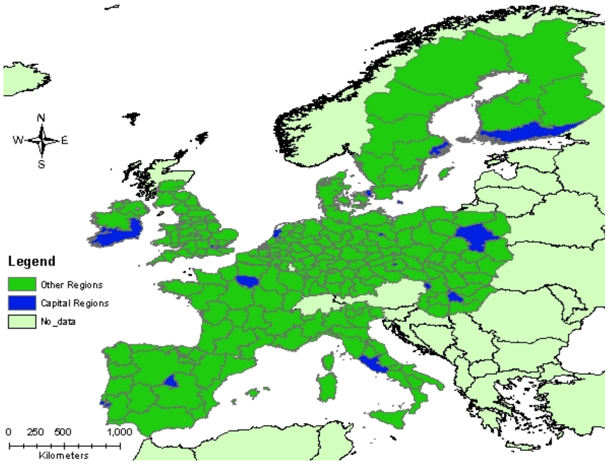


Figure 8d: Capital region regime



2.7 Descriptives and Measurement Issues

The dataset is already somewhat treated in the earlier sections where all the variables are described. Country patterns are visible in some of the dependents and independent variables. The clustering of regions of the same country, visible in Annex 1, confirms these patterns. The first small conclusions that can be made are that the regions are not homogenously divided over Europe and that it is likely that there is something going on spatially. All dependent variables however, show normal probability distributions. Employment growth is distributed with a little peak but this is not problematic. These normal distributions are also confirmed by the information the corresponding PP-plots give. Residuals are all normally distributed.

Correlations in the independents are not very high, so there is no expectation for multicollinearity. Productivity in 2000 with the wage level and Employment in 2000 with related variety are the only two sets with correlation above 0,8 but as this is below 0,9 and collinearity statistics in the regression show no alarming numbers this is not to be considered as a problem. In the absence of multicollinearity it is accepted to put all variables together in one regression model.

All independent variables were checked to investigate whether a quadratic relation suited the data better but this was never the case. Furthermore, all independents were plotted against all the dependents in search for special cases and heterogeneity but there were no special cases and heterogeneity was not found. There were no outliers found using Cook's distance. Both Irish regions scored relatively high in this test, but the scores were not exceptional so this is not alarming.

All controls are transformed into the natural logarithm. The reason for this transformation has not only to do with the normal distribution, but also with the magnitude of the figures. After taking the natural log the data is more manageable and normal distribution is better. In reaction to this transformation the dependents are also transformed into their natural log for easier interpretation. An extensive overview of all the tests discussed in this section is available upon request.

As described earlier the following countries are included in the research: Belgium, Czechoslovakia, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Slovakia, Spain, Sweden and the United Kingdom. Of course it would have been better to include countries like Switzerland, Austria, Greece etc., but due to limited data available in the Amadeus database for variety this was not possible.

Nevertheless, the countries used are an excellent sample of all European countries as almost all core European countries are included together with some new-member states so the absence of some countries is a surmountable problem.

In an ideal situation every region would have had the exact same size but unfortunately some regions are larger than other regions, which is likely to be a problem for the model. Although this size issue is partly captured in the spatial dependence there might still be significant influence of the size of the regions. The data overview gives several examples of the possible bias due to size differences. The data for Belgium is only available for NUTS-1 regions making these regions somewhat bigger than the other regions in the dataset. Although this makes the dataset weaker it is not alarming, since spatial dependence is included in the research. Rejecting the Belgian regions for their magnitude makes the dataset weaker than the data being at NUTS-1 level instead of NUTS-2.

For Sweden, data is missing for region SE07, the middle north region. For the figures used in the thesis the data of the region SE08, which is north of region SE07, is copied and used to predict the data of SE07. For the empirical analyses the region SE07 is left aside. For the United Kingdom, data is missing for the two northerly regions. These regions are left aside in the figures as well as in the empirical research. The database is further limited to regions in continental Europe, which means that French, Portuguese and Spanish overseas territories are not included.

2.7.1 Spatial Autocorrelation of the Independents

The model that eventually will be used as the most probable and best fitting is the model of equation 8. In this model the spatial dependence of the independent is taken into account. However, when looking at the independents, most of the independents seem to experience spatial dependence as well. This is tested using the Moran's I-test, where only public R&D and employment in 2000 turn out to have no spatial dependence. A full table of all the tests is shown in annex 2. All the other variables are thus subject to spatial autocorrelation. Unfortunately, it is not possible to include all the spatial dependences of these variables, due to the fact that there is multicollinearity between the different variables. Though the spatial dependence of the independent is very large and captures to a certain extent the spatial dependence of the dependents through the model, which makes the model as good as it can get.

3. Results

In this chapter, the most relevant results of the empirical research are reported. First, the two different components of output, namely employment and productivity, are discussed followed by unemployment. Moreover, for all components, the four regimes discussed in chapter 1.5 are tested.

3.1 Employment

In table 1a the outcomes of the first basic regressions are shown for employment. The first column shows the outcomes of a basic regression using ordinary least squares (OLS) without accounting for spatial dependence. More than half of the explanatory variables have significant influence on the dependent variable employment growth but as explained in the methodology this significance can be false due to the nested structure of the dataset. Therefore the spatial models (2) and (3) are introduced. The extra variable called W_Employment Growth is added to the model and in both models this variable has a highly significant positive influence. It can be concluded that the spatial models are better fit for researching employment growth. There is significant positive influence of employment growth of other regions to the own region. Another reason for using the spatial models is that the R^2, which measures the explanatory value of the total model, in models (2) and (3) is higher, where the last model shows the highest number of 0,402. This means that in model (3) 40,2% of the variety in employment growth is explained by the independent variables. The Likelihood-Ratio (LR) test for spatial lag is also highly significant indicating there is indeed a spatial dependence. Unfortunately the spatial error distribution is also significant in model (2) and almost significant in model (3), which means that more variety can be explained using other types or forms of spatial dependence. The weight matrices used capture not all the spatial dependence but it is clear that model (3) does a better job than model (2).

Table 1a: Employment Growth

Explanatory Variables	(1) OLS Model		(2) Spatial Model with Power 1		(3) Spatial Model with Power 2	
(Constant)	0,300	*0,216*	0,303	*0,189*	0,312	*0,164*
Employment 2000	-0,025	*0,013*	-0,028	*0,011*	-0,027	*0,010*
Private R&D	-0,005	*0,006*	-0,003	*0,005*	0,000	*0,005*
Public R&D	-0,006	*0,006*	-0,003	*0,005*	0,000	*0,005*
Openness Economy	0,052	*0,017*	0,048	*0,015*	0,032	*0,013*
Market Potential	-0,092	*0,021*	-0,069	*0,018*	-0,047	*0,016*
Education	0,024	*0,015*	0,020	*0,013*	0,012	*0,011*
Population Density	0,005	*0,006*	0,005	*0,005*	0,005	*0,005*
Wages	0,043	*0,011*	0,030	*0,010*	0,017	*0,009*
Related Variety	0,078	*0,040*	0,088	*0,035*	0,084	*0,030*
Unrelated Variety	0,035	*0,015*	0,029	*0,013*	0,024	*0,011*
Specialization	-0,368	*0,120*	-0,274	*0,105*	-0,194	*0,091*
W_Employment Growth			0,950	*0,035*	0,919	*0,041*
Summary Statistics:						
N	205		205		205	
R^2	0,265		0,291		0,402	
BP (heteroskedasticity)			42,002	*0,000*	45,915	*0,000*
LR (spatial lag)			37,583	*0,000*	80,966	*0,000*
LM (spatial error)			48,364	*0,000*	2,484	*0,115*

Coefficients are displayed in the first row of each model. The second row shows the standard error of these coefficients in *italics*. In the summary statistics the numbers in *italics* are probability numbers. N is the number of observations of the test. R^2 is the explanatory value of the total model. BP is the Breusch-Pagan test for heteroskedasticity, LR is the Likelihood-Ratio test for spatial dependence and LM is the Lagrange Multiplier test for the spatial error. Significance is shown by green (positive) and red (negative) colors in six gradations of alpha scores: 0,01 / 0,05 / 0,10 / 0,10 / 0,05 / 0,01

All the significant variables of model (1) maintain their significance after including the spatial variable W_Employment Growth. Some become more significant, others are less significant but they all keep their relationship with the dependent variable employment growth. Heteroskedasticity might be a problem in these models as the BP-test is significant but this can be explained by the fact that there is spatial dependence in the model. Remember that the Breusch-Pagan test can be falsely significant because this test is very sensitive to the presence of spatial dependence (Anselin, 1992). Nevertheless it is important to mention the possibility of heteroskedasticity.
Model (3) has the highest R^2 and scores substantially lower in the Lagrange-Multiplier (LM) test for spatial error than model (2) so this model suits best of the three models to explain output growth. When looking at model (3) it can be concluded that the initial employment level, employment 2000, and specialization have a negative influence on employment growth. Both types of variety are positively related to employment growth where related variety is the most significant. Of all the control variables, openness of the economy and wages are positively related to employment growth and market potential is negatively related to employment growth.

3.1.1 Regimes

In search for more explanatory power and to reduce the variety in the spatial error, regimes are tested using the same variables as in model (3) of table 1a. The regimes tested were discussed in the literature overview and the methodology earlier. An interesting extra test here is the Chow-Wald test, which demonstrates whether or not the two parts of the regime differ from each other significantly.
The first regime tested is the regime with the sizes of the cities within the regions. Table 1b gives the results for this regime, which turns out to have a reasonable effect on the explanatory power; R^2 has risen from 0,402 to 0,557 and the Chow-Wald test is also highly significant. Heteroskedasticity is a problem but the spatial error is not significant at all. The most striking observations are that unrelated variety works positive in small city regions and large city regions and specialization works negative in large city regions. Furthermore there is convergence among large city regions and divergence between small city regions.

Table 1b: Regime Size for Employment Growth

	Regime Size of Cities within Region					
Explanatory Variables	Small Cities		Medium Cities		Large Cities	
(Constant)	-0,481	*0,486*	0,904	0,544	0,727	0,243
Employment 2000	0,072	*0,039*	-0,051	0,037	-0,049	0,012
Private R&D	0,003	*0,012*	-0,009	0,008	0,003	0,006
Public R&D	-0,023	*0,012*	-0,001	0,010	0,006	0,005
Openness Economy	0,034	*0,028*	0,061	0,034	0,057	0,017
Market Potential	-0,115	*0,038*	-0,069	0,027	-0,033	0,020
Education	0,055	*0,026*	0,058	0,030	0,007	0,013
Population Density	-0,007	*0,011*	0,007	0,008	0,004	0,006
Wages	0,032	*0,027*	0,034	0,019	0,011	0,011
Related Variety	-0,042	*0,086*	0,079	0,055	0,070	0,038
Unrelated Variety	0,063	*0,028*	-0,043	0,019	0,044	0,015
Specialization	0,048	*0,168*	-0,194	0,258	-0,377	0,119
W_Employment Growth	0,892	*0,050*				
Summary Statistics:						
N			205			
R^2			0,557			
Chow-Wald			62,4	0,000		
BP (heteroskedasticity)			9,501	0,009		
LR (spatial lag)			77,2	0,000		
LM (spatial error)			0,105	0,746		

See notes below table 1a. Coefficients that differ significantly with an alpha of 0,05 are boxed.

Table 1c shows the results of the three other regimes for employment growth. All regimes show higher explanatory power than model (3) in table 1a but the regime distinguishing between objective-1 regions and other regions stands out with an R^2 of 0,527 although this is a lower score compared to the size of cities regime. This objective-1 regime is also the only regime of the three where the Chow-Wald test has a significant outcome and it has the best score for the spatial error dependence. Heteroskedasticity is again a possible problem in all models. An interesting observation though is that the coefficient of convergence is a little larger in non-capital regions compared to the total dataset (-0,038 instead of –0,027). As the Chow-Wald test is not significant for the top university regime and the capital regime no further conclusions will be made on these regimes. Being an objective-1 region means that significantly different independent variables have an influence on the employment growth compared to the other regions. In objective-1 regions public R&D has a negative influence on employment growth. Education, the wage level and related variety are positively related to employment growth. Unrelated variety has no positive influence anymore in objective-1 regions.
Individual Chow-Wald tests indicate that public R&D and the wage level in regime (1) are the only independents that differ from each other significantly with an alpha of 0,05. The spatial variable W_Employment Growth is still highly significant.

Table 1c: Other Regimes for Employment Growth

	(1) Regime Objective-1				(2) Regime Top 75 University				(3) Regime Capital Region			
Explanatory Variables	No Obj 1		Obj1		No Top University		Top University		No Capital		Capital	
(Constant)	0,637	*0,244*	0,093	*0,281*	0,519	*0,215*	0,687	*0,426*	0,706	*0,212*	-0,146	*0,536*
Employment 2000	-0,020	*0,012*	-0,028	*0,017*	-0,032	*0,011*	0,018	*0,030*	-0,038	*0,011*	0,030	*0,046*
Private R&D	-0,004	*0,006*	0,000	*0,008*	-0,001	*0,005*	-0,008	*0,011*	0,000	*0,005*	-0,010	*0,025*
Public R&D	0,006	*0,005*	-0,040	*0,011*	-0,001	*0,005*	0,004	*0,014*	0,002	*0,005*	-0,017	*0,017*
Openness Economy	0,021	*0,023*	0,030	*0,018*	0,032	*0,015*	0,000	*0,042*	0,024	*0,014*	0,065	*0,063*
Market Potential	-0,049	*0,023*	-0,076	*0,025*	-0,049	*0,017*	-0,065	*0,038*	-0,050	*0,017*	-0,043	*0,060*
Education	0,003	*0,014*	0,040	*0,019*	0,015	*0,011*	-0,084	*0,058*	0,010	*0,011*	0,035	*0,070*
Population Density	0,007	*0,006*	0,008	*0,008*	0,009	*0,005*	-0,005	*0,010*	0,008	*0,005*	-0,019	*0,021*
Wages	-0,010	*0,012*	0,065	*0,014*	0,019	*0,010*	-0,004	*0,023*	0,014	*0,009*	0,008	*0,033*
Related Variety	0,062	*0,036*	0,186	*0,056*	0,096	*0,035*	0,069	*0,071*	0,108	*0,033*	-0,069	*0,120*
Unrelated Variety	0,026	*0,013*	0,000	*0,020*	0,033	*0,013*	0,002	*0,022*	0,018	*0,011*	0,068	*0,084*
Specialization	-0,103	*0,165*	-0,137	*0,114*	-0,202	*0,101*	0,184	*0,293*	-0,244	*0,096*	-0,217	*0,370*
W_Employment Growth	0,865	*0,059*			0,920	*0,041*			0,924	*0,039*		
Summary Statistics:												
N	205				205				205			
R^2	0,527				0,437				0,447			
Chow-Wald	40,2	*0,000*			13,2	*0,355*			18,6	*0,100*		
BP (heteroskedasticity)	18,958	*0,000*			12,812	*0,000*			3,565	*0,059*		
LR (spatial lag)	65,9	*0,000*			79,6	*0,000*			83,4	*0,000*		
LM (spatial error)	0,457	*0,499*			2,072	*0,150*			1,042	*0,307*		

See notes below table 1a. Coefficients that differ significantly with an alpha of 0,05 are boxed.

3.2 Productivity

Employment growth can foster economic growth but when employment doesn't change productivity growth of the existing employees is another way to attain economic growth. Here, the same three simple regressions are executed as for output growth and employment growth. The results are presented in table 2a. Model (3) is again the best fit, with an R^2 of 0,837. This means that 83,7% of the variety in productivity growth is explained by all the independents together in the model. Heteroskedasticity continues to be a problem but interesting is that in model (3) the spatial error test is no longer significant. Spatial dependence is captured quite well with the introduction of W_Productivity growth. The initial productivity level has a negative influence on productivity growth and specialization has a positive effect. There is no significant influence of any type of variety visible for productivity growth. The control variables private R&D, market potential and education are positively related to productivity growth and openness of the economy is negatively related to productivity growth.

Table 2a: Productivity Growth

Explanatory Variables	(1) OLS Model		(2) Spatial Model with Power 1		(3) Spatial Model with Power 2	
(Constant)	0,463	*0,167*	0,048	*0,140*	-0,036	*0,121*
Productivity 2000	-0,223	*0,018*	-0,168	*0,015*	-0,092	*0,014*
Private R&D	0,024	*0,005*	0,017	*0,004*	0,007	*0,004*
Public R&D	-0,006	*0,005*	-0,005	*0,004*	-0,001	*0,004*
Openness Economy	-0,032	*0,016*	-0,025	*0,013*	-0,022	*0,011*
Market Potential	0,079	*0,018*	0,075	*0,015*	0,055	*0,013*
Education	0,066	*0,012*	0,054	*0,010*	0,030	*0,009*
Population Density	-0,003	*0,005*	-0,004	*0,004*	-0,004	*0,004*
Wages	0,007	*0,014*	0,008	*0,012*	0,004	*0,010*
Related Variety	-0,028	*0,021*	-0,026	*0,018*	-0,018	*0,015*
Unrelated Variety	0,005	*0,013*	0,004	*0,011*	0,000	*0,009*
Specialization	0,560	*0,103*	0,414	*0,086*	0,268	*0,074*
W_Productivity Growth			0,960	*0,027*	0,899	*0,043*
Summary Statistics:						
N	205		205		205	
R^2	0,759		0,781		0,837	
BP (heteroskedasticity)			55,010	*0,000*	78,453	*0,000*
LR (spatial lag)			57,061	*0,000*	113,714	*0,000*
LM (spatial error)			39,490	*0,000*	0,051	*0,821*

See the notes below table 1a.

3.2.1 Regimes

Just like in employment growth the four regimes are also tested for productivity growth. Table 2b shows the results of the first test. Just like for employment, the size of cities regime is significant. 86,1% of all variation in productivity growth is explained in this test. Heteroskedasticity is still present just like the spatial dependence and the spatial error is not significant. Specialization only works in small city regions and large city regions. Population density has a surprising negative effect in small city regions.
Table 2c shows the other regimes. For all regimes there is a significant difference visible in the two sides of the regimes as the Chow-Wald tests are significant in all models. Notable is that heteroskedasticity is not a problem anymore in model (2) and (3). The spatial variable W_Productivity Growth continuous to be highly significant. The initial level of productivity has no negative influence in objective-1 regions and top university regions indicating the absence of convergence. Another interesting detection is that the coefficient for convergence is lower in both non-capital and capital regions compared to the total model. There is more convergence when looking at the two sides of the capital regimes separately than when looking at the total dataset.

The separation thus has a positive influence on convergence, which indicates that there is a diverging force at stake. Specialization is in particular of positive influence in other regions than top university-regions and it works significantly better in capital regions than in non-capital regions as indicated by the significance of the individual Chow-Wald test. Unrelated variety works positive in objective-1 regions in contrast to other regions.

Table 2b: Regime Size for Productivity Growth

	Regime Size of Cities within Region					
Explanatory Variables	Small Cities		Medium Cities		Large Cities	
(Constant)	-0,791	*0,257*	-0,210	0,237	-0,211	0,179
Productivity 2000	-0,107	*0,033*	-0,122	0,027	-0,072	0,017
Private R&D	-0,020	*0,010*	0,018	0,007	0,004	0,005
Public R&D	0,029	*0,010*	-0,004	0,008	-0,005	0,004
Openness Economy	0,033	*0,021*	-0,035	0,028	-0,036	0,015
Market Potential	0,091	*0,032*	0,037	0,022	0,042	0,017
Education	0,064	*0,019*	0,024	0,024	0,021	0,010
Population Density	-0,029	*0,008*	-0,001	0,006	0,004	0,005
Wages	0,029	*0,026*	0,023	0,023	-0,001	0,012
Related Variety	-0,043	*0,051*	-0,019	0,041	-0,009	0,025
Unrelated Variety	0,011	*0,026*	-0,001	0,015	-0,012	0,012
Specialization	0,376	*0,138*	0,244	0,212	0,273	0,096
W_Productivity Growth	0,932	*0,033*				
Summary Statistics:						
N			205			
R^2			0,861			
Chow-Wald			46,3	0,004		
BP (heteroskedasticity)			7,157	0,028		
LR (spatial lag)			126,5	0,000		
LM (spatial error)			1,287	0,257		

See notes below table 1a. Coefficients that differ significantly with an alpha of 0,05 are boxed.

Table 2c: Other Regimes for Productivity Growth

	(1) Regime Objective-1				(2) Regime Top 75 University				(3) Regime Capital Region			
Explanatory Variables	No Obj 1		Obj1		No Top University		Top University		No Capital		Capital	
(Constant)	0,033	*0,190*	0,232	*0,191*	-0,024	*0,139*	-1,071	*0,302*	-0,230	*0,123*	-1,156	*0,392*
Productivity 2000	-0,127	*0,027*	0,017	*0,029*	-0,098	*0,016*	0,103	*0,075*	-0,102	*0,013*	-0,153	*0,058*
Private R&D	0,003	*0,005*	0,022	*0,006*	0,015	*0,004*	-0,019	*0,009*	0,010	*0,003*	-0,029	*0,025*
Public R&D	0,003	*0,004*	-0,020	*0,009*	0,000	*0,004*	-0,010	*0,011*	-0,004	*0,003*	0,035	*0,012*
Openness Economy	-0,047	*0,020*	-0,020	*0,014*	-0,040	*0,012*	-0,021	*0,034*	-0,041	*0,010*	0,035	*0,036*
Market Potential	0,051	*0,019*	0,086	*0,019*	0,044	*0,013*	0,065	*0,029*	0,048	*0,012*	0,120	*0,033*
Education	0,050	*0,011*	-0,010	*0,015*	0,024	*0,009*	-0,022	*0,044*	0,026	*0,008*	0,099	*0,050*
Population Density	-0,002	*0,005*	-0,019	*0,006*	-0,005	*0,004*	-0,011	*0,008*	-0,002	*0,004*	-0,037	*0,013*
Wages	0,009	*0,010*	-0,096	*0,027*	-0,006	*0,011*	0,011	*0,018*	0,009	*0,009*	0,026	*0,049*
Related Variety	-0,020	*0,016*	-0,007	*0,027*	-0,028	*0,016*	-0,015	*0,034*	-0,025	*0,014*	-0,095	*0,079*
Unrelated Variety	-0,009	*0,010*	0,042	*0,015*	-0,010	*0,010*	0,017	*0,017*	-0,010	*0,008*	0,094	*0,071*
Specialization	0,268	*0,146*	0,156	*0,088*	0,153	*0,077*	0,334	*0,296*	0,188	*0,068*	0,965	*0,260*
W_Productivity Growth	0,937	*0,032*			0,895	*0,044*			0,890	*0,043*		
Summary Statistics:												
N		205				205				205		
R^2		0,861				0,864				0,887		
Chow-Wald		49,6	*0,000*			39,3	*0,000*			85,8	*0,000*	
BP (heteroskedasticity)		5,520	*0,019*			0,270	*0,603*			0,033	*0,857*	
LR (spatial lag)		129,3	*0,000*			115,5	*0,000*			127,6	*0,000*	
LM (spatial error)		2,885	*0,089*			0,222	*0,638*			0,742	*0,389*	

See notes below table 1a. Coefficients that differ significantly with an alpha of 0,05 are boxed.

3.3 Unemployment

The last dependent variable that is tested is the unemployment growth of the regions. A negative influence here indicates that unemployment growth is negatively related to the variable, which of course means that unemployment growth is lower when this variable increases. Negative relations are here thus positive for the regional economic performance. Table 3a shows the simple regressions where again model (3) is the best fit. Now 81,4% of the variety is explained and the spatial error test is not significant. Heteroskedasticity is again a possible problem though.

The initial unemployment has a negative influence on unemployment growth. The unemployment growth of neighbor regions is not good for unemployment growth of the own region as it has a strong positive influence. The only other variables with a positive influence are the openness of the economy and related variety. Further, private R&D, market potential, education and specialization all have a negative influence on unemployment growth.

Table 3a: Unemployment Growth

Explanatory Variables	(1) OLS Model		(2) Spatial Model with Power 1		(3) Spatial Model with Power 2	
(Constant)	1,289	*1,069*	1,536	*0,847*	1,393	*0,788*
Unemployment 2003	-0,561	*0,034*	-0,453	*0,027*	-0,371	*0,030*
Private R&D	-0,105	*0,035*	-0,076	*0,028*	-0,053	*0,026*
Public R&D	-0,032	*0,033*	0,008	*0,026*	0,008	*0,025*
Openness Economy	0,500	*0,099*	0,517	*0,078*	0,428	*0,073*
Market Potential	-0,592	*0,114*	-0,430	*0,090*	-0,324	*0,086*
Education	-0,064	*0,081*	-0,129	*0,064*	-0,118	*0,060*
Population Density	0,083	*0,033*	0,042	*0,026*	0,034	*0,025*
Wages	0,225	*0,065*	0,124	*0,052*	0,078	*0,049*
Related Variety	0,138	*0,137*	0,206	*0,109*	0,191	*0,102*
Unrelated Variety	0,134	*0,084*	0,051	*0,067*	-0,006	*0,063*
Specialization	-4,238	*0,666*	-3,101	*0,528*	-2,305	*0,501*
W_Unemployment Growth			0,961	*0,027*	0,758	*0,058*
Summary Statistics:						
N	205		205		205	
R^2	0,717		0,766		0,814	
BP (heteroskedasticity)			24,087	*0,012*	23,267	*0,016*
LR (spatial lag)			77,163	*0,000*	101,389	*0,000*
LM (spatial error)			21,683	*0,000*	1,992	*0,158*

See the notes below table 1a.

3.3.1 Regimes

The four regimes used in employment growth and productivity growth, are also applied on unemployment growth. Table 3b shows the results of the size regime. For the third time this regime is significant. The summary statistics show the same picture as for the other dependents. Specialization works negative in small city regions and large city regions.

In table 3c the other regimes are presented. It is notable that all models have a significant Chow-Wald test and heteroskedasticity is not a potential problem in any of the three models. Only model (3) shows a significant spatial error dependence, which makes the model less good.

The spatial variable W_Unemployment Growth is positively related to unemployment growth. Furthermore it is interesting to see that specialization doesn't have its negative influence in none objective-1 regions and regions containing a university. For related variety there are no highly significant influences visible. Unrelated variety has a positive influence in capital-regions and related variety has a negative influence in regions not containing a capital city.

Unrelated variety can be concluded to work negatively for economic performance through unemployment growth. Convergence is not visible in capital regions.

Table 3b: Regime Size for Unemployment Growth

Explanatory Variables	Small Cities		Medium Cities		Large Cities	
(Constant)	4,124	*1,572*	-1,548	1,615	2,087	1,118
Unemployment 2003	-0,386	*0,064*	-0,310	0,061	-0,356	0,033
Private R&D	0,042	*0,073*	-0,014	0,048	-0,060	0,034
Public R&D	-0,100	*0,071*	0,067	0,053	0,004	0,028
Openness Economy	0,572	*0,158*	0,138	0,192	0,466	0,098
Market Potential	-0,703	*0,216*	0,006	0,155	-0,371	0,115
Education	-0,230	*0,134*	-0,347	0,173	-0,087	0,071
Population Density	0,191	*0,059*	-0,052	0,038	0,021	0,036
Wages	0,091	*0,136*	0,160	0,120	0,072	0,063
Related Variety	-0,157	*0,310*	0,227	0,283	0,127	0,172
Unrelated Variety	0,227	*0,159*	0,020	0,109	-0,049	0,084
Specialization	-2,987	*0,948*	-0,517	1,511	-2,589	0,674
W_Unemployment Growth	0,802	*0,054*				
Summary Statistics:						
N			205			
R^2			0,839			
Chow-Wald			36,4	0,050		
BP (heteroskedasticity)			8,305	0,016		
LR (spatial lag)			109,1	0,000		
LM (spatial error)			0,019	0,890		

See notes below table 1a. Coefficients that differ significantly with an alpha of 0,05 are boxed.

Table 3c: Other Regimes for Unemployment Growth

	(1) Regime Objective-1				(2) Regime Top 75 University				(3) Regime Capital Region			
Explanatory Variables	No Obj 1		Obj1		No Top University		Top University		No Capital		Capital	
(Constant)	-0,142	*1,212*	0,613	*1,253*	0,546	*0,876*	4,793	*2,063*	2,517	*0,840*	-1,119	*2,120*
Unemployment 2003	-0,346	*0,038*	-0,410	*0,058*	-0,382	*0,030*	-0,286	*0,073*	-0,385	*0,029*	-0,072	*0,103*
Private R&D	-0,040	*0,033*	-0,058	*0,044*	-0,078	*0,028*	-0,024	*0,062*	-0,047	*0,025*	0,494	*0,138*
Public R&D	0,014	*0,026*	-0,038	*0,065*	0,015	*0,027*	0,061	*0,075*	0,031	*0,024*	-0,124	*0,090*
Openness Economy	-0,080	*0,123*	0,586	*0,117*	0,497	*0,078*	-0,116	*0,222*	0,413	*0,071*	1,039	*0,282*
Market Potential	-0,089	*0,121*	-0,377	*0,138*	-0,293	*0,093*	-0,315	*0,176*	-0,367	*0,090*	-0,564	*0,257*
Education	-0,051	*0,076*	-0,086	*0,104*	-0,116	*0,059*	-0,337	*0,314*	-0,164	*0,058*	-0,426	*0,346*
Population Density	-0,006	*0,031*	0,089	*0,045*	0,055	*0,028*	0,006	*0,054*	0,012	*0,027*	0,259	*0,094*
Wages	-0,038	*0,067*	0,178	*0,082*	0,132	*0,052*	-0,238	*0,123*	0,027	*0,048*	0,447	*0,171*
Related Variety	0,228	*0,127*	0,047	*0,189*	0,159	*0,110*	0,469	*0,276*	0,229	*0,101*	-0,476	*0,480*
Unrelated Variety	0,055	*0,078*	-0,012	*0,108*	-0,036	*0,071*	0,132	*0,126*	0,003	*0,059*	0,936	*0,440*
Specialization	0,105	*0,933*	-2,606	*0,628*	-2,232	*0,536*	1,183	*1,544*	-2,953	*0,502*	-3,224	*1,928*
W_Unemployment Growth	0,714	*0,063*			0,784	*0,058*			0,767	*0,056*		
Summary Statistics:												
N		205				205				205		
R^2		0,840				0,832				0,844		
Chow-Wald		30,1	*0,003*			24,7	*0,016*			40,6	*0,000*	
BP (heteroskedasticity)		0,243	*0,622*			2,099	*0,147*			0,016	*0,899*	
LR (spatial lag)		86,1	*0,000*			108,7	*0,000*			107,3	*0,000*	
LM (spatial error)		0,026	*0,872*			1,005	*0,316*			5,239	*0,022*	

See notes below table 1a. Coefficients that differ significantly with an alpha of 0,05 are boxed.

4. Discussion

In this part the outcomes of the method section are discussed, using the literature review, the research model and the results of the empirical research.

In every single test there is spatial dependence of the dependent variable. All the Likelihood-Ratio tests are highly significant, just like every spatial lag in all tests. Even before the empirical research, looking at the spatial patterns visible in some of the variables already made this conclusion. Models not accounting for spatial dependence can be considered incompetent and incorrect. European regions are thus definitely dependent upon each other, so the first hypothesis can be fully accepted. Because all regions are dependent upon each other, regional policy makers should not only focus on the own region but should also take into account the effects new policy has on other regions. Furthermore, policy makers should look at existing and new policies in neighbor regions and investigate the effects of these policies on the own region.

Convergence between regions is visible in almost every test so despite the fact that hypothesis two cannot be accepted as convergence is not visible in every test, there are very strong indications for convergence. The only exception is in small city regions where divergence is visible for employment growth.

Another interesting finding related to the convergence-divergence debate is that when capital regions are left aside, convergence is stronger in all the dependents. The capital regime test for productivity growth shows that the division of capital regions and non-capital regions results in more convergence in both groups. The division of capital regions and non-capital regions clearly influences the convergence. Actually, without this regime in the total dataset there is less convergence than measured in both groups of the regimes separately. It can be concluded that a diverging force that is related to the capital regions regime works here negatively on the convergence.

Education has a positive influence on growth through higher productivity and lower unemployment. It has a smaller positive influence on employment growth. The hypothesis on education can thus be accepted and it can be concluded that education is positively related to economic growth as the hypothesis suggests. Investing in education is thus a good choice in order to attain economic growth. Policy makers should continue to invest in human capital, even in tougher economic times like now, because this is one of the easiest ways to reach economic growth.

The results for the different sector structure variables support the choice for more in-depth research in the specialization-diversity debate because there are surely differences visible for the influence of related variety, unrelated variety and specialization on the employment growth, productivity growth and unemployment growth.

Specialization has a negative effect on employment growth and a positive effect on productivity growth. The relationships though are not robust. Specialization works exceptionally well for productivity growth in capital regions, it works in non-capital regions but it doesn't work in regions containing a medium-sized city or top university regions. Unlike expected, there is no difference in the influence between objective-1 and non-objective-1 regions. The third hypothesis can thus not be accepted although there are strong indications that specialization does work for productivity growth.

Both types of variety positively influence employment growth. This is in line with the findings of Frenken et al. (2007) who found that for Dutch urban regions related variety is positively related to the beneficial knowledge spillovers and unrelated variety is found to be better against economic shocks. The same results are found in Britain (Bishop & Gripaios, 2010) and Italy (Boschma and Iammarino, 2009). Related variety in the EU-wide case turns out to be more valuable though for employment growth than unrelated variety. When looking at the different regimes, related variety is not in every case of positive influence on employment growth; it only works in objective-1 regions, regions without a top university and non-capital regions. Variety has no positive influence on productivity growth, except for the positive influence of unrelated variety when looking at objective-1 regions. Related variety has a marginal significant positive influence on unemployment growth instead of the expected negative influence. Unrelated variety is surprisingly positively related to unemployment growth only in capital regions.

It can be concluded that hypothesis four cannot be accepted. The first part (part a) holds, but is not robust. The second part where unrelated variety is also expected to be strongly negative on unemployment growth needs to be rejected as well.
Especially for unemployment growth, results of the tests are not as expected. This may be the result of the measurement problems in measuring unemployment. Unemployment might be affected by discouraged workers that are willing to work but stopped looking for work as a result of not finding a job, which means that these discouraged workers are not accounted as unemployed anymore. Another measurement problem has to do with part-time workers as they don't work full time but are not indicated unemployed for the part they don't work. More part-time workers thus result in lower unemployment figures.
Taking into account the different results found for related and unrelated variety policy makers can now focus better on a specific type of variety to reach their goals instead of just supporting diversity in general. Variety, specifically related variety, works for most regions that lag behind in employment levels. Specialization works for most regions that want to increase their productivity. With the inclusion of the different regimes policy can be set out even more specific to reach the desired targets.
In most cases related variety is more positively related to growth than unrelated variety. It is thus better to support related variety rather than unrelated variety. Firms operating in the same sector should be brought in contact with each other, even beyond regional boundaries. In an ideal situation, all firms operating in the same sector are clustered in one place to maximize the benefits of the related variety.
Out of the twelve regimes tested, ten turned out to be of significant influence. The only two non-significant regimes are the top university regime and the capital regime for employment growth. The last hypothesis can be accepted partially and it can be concluded that regional performance is heterogeneously spread over Europe. Every region should first consider to which regime it belongs and where their opportunities lie before implementing policy on specific aspects to reach economic growth.
The objective-1 regime is even always significant, indicating that objective-1 regions should always be treated differently to attain economic growth compared to non-objective-1 regions. The denotation of objective-1 regions is hereby supported as the two groups respond significantly different to certain policies. Cohesion policy can now focus on those factors that specifically higher economic growth in objective-1 regions.

Conclusion

Regional economic performance in Europe is a complex topic to investigate. The influence of both related and unrelated variety and also of specialization is quite visible. Specifically related variety seems to go hand in hand with employment growth and specialization links to productivity growth. The relationships however are not robust; when distinguishing between different regimes the results are not always as expected. There is still improvement possible in the measurement method for related variety because it cannot be said with certainty that the method used in this paper comes to the best possible figures for related variety. Other methods might result in other conclusions.

Unrelated variety does not work for reducing unemployment growth like expected by portfolio theory, possibly due to the complex way that unemployment is measured.

Convergence is widely present in Europe, although capital regions contribute negatively to this convergence presumably due to centralization policy. The expected convergence in the neoclassical growth theory and the new growth theory is definitely found. Economic growth is dependent on initial conditions and interregional interactions, where regions lagging behind can benefit from spillovers of neighbor-regions.

Country patterns are also visible in this convergence issue. Regions of the same country tend to cluster in their initial economic performance and their growth rates. Actually, all spatial tests show that there is definitely something going on spatially. In the future, spatial dependence should never be neglected when researching European regions. More detailed research might even tell us more about this spatially dependence between regions. A suggestion is to use models with fixed effects or multilevel analysis, in which country effects can be captured more accurately. Furthermore, other matrices than geographical distance can be used to measure spatial dependence like network distance or travel distance for example. Even though travel distance in particular is very similar to geographical distance.

The results are possibly biased due to the measurement issues discussed earlier. The two most important issues are the almost constantly returning presence of heteroskedasticity according to the significance of the Breusch-Pagan test and the differences in size of the regions. Heteroskedasticity might not be a problem because the performed Breusch-Pagan test doesn't account for the existence of spatial dependence but this cannot be said with certainty. Adding spatial dependence and the control variable population density account for the size of the regions but it can still be problematic for the model. The significance of the spatial error in the Lagrange Multiplier test in some of the models supports this possible bias.

Through the significance of the regimes, heterogeneity is proven like expected in the new economic geography theory, thereby supporting region-specific policy strategies. Higher levels of economic performance in larger, densely populated areas also reflect the heterogeneity, which is in line with urbanization theory.

Regions with medium-sized cities show no significant positive influence of the sector structure variables on the dependent variable. Apparently it does not matter for these regions whether the region has more related variety, unrelated variety or specialization. Unrelated variety in medium-sized city regions even has a negative influence on employment growth. However, regions with small cities or large cities do benefit from unrelated variety for employment growth and specialization for productivity growth. Specialization in these regions is also good to reduce unemployment. A mix of unrelated variety and specialization works best in these regions, while for medium-sized city regions it has no influence. Regions are recommended to use a region specific policy mix to achieve maximum growth. These outcomes speak against the expectation that focus should be on medium-sized city regions like Barca et al. (2012) suggested while specialization in the best practices is supported for regions with small cities or large cities. The polycentric structure of medium-sized city regions is not proven to be beneficial for economic growth.

The recommended mix of variety and specialization means that no choice can be made in whether localization externalities or Jacobs externalities works best for economic growth. It is clear that they both have positive influence, variety through employment growth and specialization through productivity growth.

Investing in knowledge is always a good decision, specifically when the region lags behind in productivity or experiences high unemployment figures. Region-specific people-based policy focused on education is thus recommended. This investment in knowledge may possibly also be beneficial for other regions due to spillovers just like is stated in the new growth theory and the EEG. Agreements with neighbour regions to invest in knowledge together might be a good alternative here. Upholding good relationships with neighbour regions and keeping track of their development is generally a smart approach, since spatial dependence cannot be neglected in regional economic growth. Further research should focus even more on the spatial dependence of the European regions and better measurement methods for related variety.

References

Acs, Z. J., Braunerhjelm, P., Audretch, D. B. & Carlsson, B. (2009). The Knowledge Spillover Theory of Entrepreneurship. *Small Business Economics, 32(1),* 15-30.

Anselin, L. (1992). *Spacestat tutorial. A Workbook for Using SpaceStat in the Analysis of Spatial Data.* Urbana: University of Illinois

Attaran, M. (1986). Industry Diversity and Economic Performance in U.S. Areas. *Annals of Regional Science, 20,* 44-54.

Audretsch, D. B. & Feldman, M., P. (1995). *Innovative Clusters and the Industry Life Cycle.* CEPR Discussion Papers no. 1161. Centre for Economic Policy Research, London.

Barca, F. (2009). *An agenda for a Reformed Cohesion Policy: a Place-based Approach to Meeting European Union Challenges and Expectations*. Independent Report prepared at the request of Danuta Hübner, Commissioner for Regional Policy, Brussels.

Barca, F., McCann P. & Rodriguez-Pose, A. (2012). The case for regional development intervention: Place-Based versus Place-Neutral Approaches. *Journal of Regional Science, 52,* 134–152.

Beaudry, C. & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy, 38,* 318-337.

Bishop, P. and P. Gripaios (2010), 'Personal insolvency in England and Wales: a spatial analysis'. *Urban Studies 47(8)*: 1687-1702.

Boschma, R. A. & Lambooy, J. G. (1999). Evolutionary Economics and Economic Geography. *Journal of Evolutionary Economics, 9(4),* 411-429.

Boschma, R. A. & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography 85(3)*: 289-311.

Bosma, N. S. & Van Oort, F. (2013). Agglomeration Economies, Inventors and Entrepreneurs as Engines of European Regional Productivity. *The Annals of Regional Science*. Forthcoming.

Brulhart, M. & Mathys, N. A., (2008). Sectoral agglomeration economies in a panel of European regions. *Regional Science and Urban Economics, 38,* 348-362.

Caniels, M. C. J. (2002) *Knowledge Spillovers and Economic Growth – Regional Growth Differentials across Europe*. Presented at the BABES colloquium.

Combes, P. P., Mayer, T. & Thisse, J.F. (2008)*, Economic geography. The integration of regions and nations*. Princeton: Princeton University Press.

Cornett, A. P., & Sorensen, N. K. (2010). Regional Divergence versus European Convergence with Special Emphasis on the Northern Hemisphere. *Public Administration & Regional Studies, 2(6),* 7-36.

Cortright, J. (2001). New Growth Theory, Technology and Learning. *Reviews of Economic Development Literature and Practice, 4,* 2-35.

David, P. A. (1985). Clio and the Economics of QWERTY. *American Economic Review, 75(2),* 332-37.

Diamond, C. A., & Simon, C. J. (1990) Industrial Specialization and the Returns to Labor. *Journal of Labor Economics, 8(2),* 175–201**.**

Dogaru, T., Van Oort, F., & Thissen, M. (2011). Agglomeration Economies in European Regions: Perspectives for Objective 1 Regions. *Tijdschrift voor Economische en Sociale Geografie, 102(4),* 486-494.

Ellison, G., Glaeser, E. L., & Kerr, W. (2007). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. Working Paper 13068, National Bureau of Economic Research.

European Union (2011). *Connecting Universities to Regional Growth: A Practical Guide* Newcastle: J. Goddard.

Fujita, M., Krugman, P. R. & Venables, A. J. (1999), *The Spatial Economy: Cities, Regions and International Trade.* Cambridge, Massachusetts: MIT Press.

Frenken, K. & Boschma R. A. (2007). Theoratical Framework for EEG: Industrial Dynamics and Urban Growth as a Branching Process. *Journal of Economic Geography, 7(5),* 635-649.
Frenken, K. & Hoekman, J. (2006). Convergence in an enlarged Europe: the role of network cities. *Journal of Economic and Social Geography, 97(3),* 321-326.
Frenken, K., Van Oort, F. G. & Verburg T. (2007). Related Variaty, Unrelated Variety and Regional Economic Growth. *Regional Studies, 41(5),* 685-697.
Friedmann, J. (1986). The World City Hypothesis. *Development and Change, 17(1),* 69 83.
Gardiner, B., Martin R., & Tyler P., (2004). Competitiveness, productivity and economic growth across the European regions. *Regional Studies, 38(9),* 1045-1068.
Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. & Schleifer, A. (1992) Growth in Cities. *Journal of Political Economy, 100(6),* 1126-1152.
Haug, P. (2004). Diversifikation und Regionale Wirtschafts- und Beschäftigungsentwicklung – Eine empirische Analyse Für Ausgewählte Deutsche Gebiete. *Jahrbuch für Regionalwissenschaft, 24(2),* 177-195.
Jacobs, J. (1969). *The Economy of Cities.* New York: Random House.
Jaffe, A. B., Trajtenberg, M. & Henderson R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, *108,* 577-598.
Kaldor, N. (1970). The Case for Regional Policies. *Scottish Journal of Political Economy, 17,* 337-348.
Kim, S. & Law, M. (2012). Histories, Institutions and Cities: a view from the Americas. *Journal of Regional Science, 52(1),* 10-39.
Krugman, P. R. (1979). Increasing Returns, Monopolistic Competion and International Trade. *Journal of International Economics, 9,* 469-479.
Krugman, P. R. (1991a). Increasing Returns and Economic Geography. *Journal of Political Economy*, *99(3),* 483-499.
Krugman, P. R. (1991b). *Geography and Trade.* Cambridge, Massachusetts: MIT Press.
Lancaster, K. (1990). The Economics of Product Variety: a survey. *Marketing Science, 9(3),* 189-206.
LeSage, J. P. & Pace R. K. (2009). *Introduction to Spatial Economics.* Boca Raton, Florida: Taylor & Francis.
Markowitz, H. M. (1959). *Porfolio Selection: Efficient Diversification of Investments.* New York: John Wiley & Sons.
Marrocu, E., Paci, R., & Usai, S. (2012). Productivity Growth in the Old and New Europe: the Role of Agglomeration Externalities. *Journal of Regional Science.* Early view.
Marshall, A. (1890). *Principles of Economics.* London: MacMillan.
Martin, R. & Sunley, P. (1998). Slow convergence? The new Endogenous Growth Theory and Regional Development. *Economic Geography, 74(3),* 201–227.
McCann, P. & Ortega-Argilés, R. (2011). *Smart Specialization, Regional Growth and Applications to EU Cohesion Policy.* Working papers 2011/14, Institut d'Economia de Barcelona (IEB).
Montgomery, C. (1994). Corporate diversification. *Journal of Economic Perspectives, 8,* 163-178.
Moreno, R., Paci R. & Usai S. (2005). Spatial spillovers and innovation activity in European regions. *Environment & Planning A, 37,* 1793-1812.
Myrdal, G. (1957). *Economic Theory and Under-developed Regions.* London: Gerald Duckworth.
OECD (2007). *Higher Education and Regions; Globally Competitive, Locally Engaged,* OECD publishing.
PBL (2010). *Measuring regional competitiveness in Europe using a new detailed IO-table*. The Hague: PBL.
PBL (2011). *The determinants of FDI and economic growth in European regions*. The Hague: PBL.

Peng, X., Chen, X., & Cheng, Y. (2010). *Urbanization and Its Consequences.* In Demography, Vol. II, UNESCO-The Encyclopedia of Life Support Systems (EOLSS) Joint Committee. New York: United Nations.

Petrakos, G., Kallioras, D., & Anagnostou, A. (2007). *A Generalized Model of Regional Economic Growth in the European Union,* DYNREG Research Project – Dynamic Regions in a Knowledge Driven Global Economy: Lessons and Policy Implications for the European Union, Economic and Social Research Institute (ESRI), WP

Porter, M. (1990), *The competitive advantage of nations.* New York: Free Press.

Porter, M. (2003), The economic performance of regions. *Regional Studies, 37*, 549-578.

Puga, D. (2002). European Regional Policies in Light of Recent Location Theories. *Journal of Economic Geography, 2,* 373–406.

Quigley, J. M. (2008). Urban Economics*.*. In S. N. Durlauf & L. E. Blume, *The New Palgrave Dictionary of Economics*. New York: Palgrave Macmillan.

Romer, P. M. (1986). Increasing Returns and Long-Run Growth, *The Journal of Political Economy, 94(5),* 1002-1037.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy, 98(5),* 71-102.

Rosenthal, S. S. & Strange, W. C. (2001). The Determinants of Agglomeration. *Journal of Urban Economics, 50(2),* 191-229.

Sassen, S., (2001). *The Global City: New York, London, Tokyo.* New Jersey: Princeton University Press.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics, 70(1),* 65-94.

Van Oort, F.G. (2004), *Urban growth and innovation. Localized externalities in the Netherlands*. Aldershot: Ashgate.

Van Oort, F. G. (2013). Agglomeration Economics Beyond the Specialization-Diversity Controversy. In Karlsson, C. & Andersson, M. *Handbook of Research Methods and Applications in Economic Geography.* Berlin: Springer. Forthcoming.

Von Thünen, J. H. (1966). *Isolated state: an English edition of Der Isolierte Staat.* Oxford, New York: Pergamon Press.

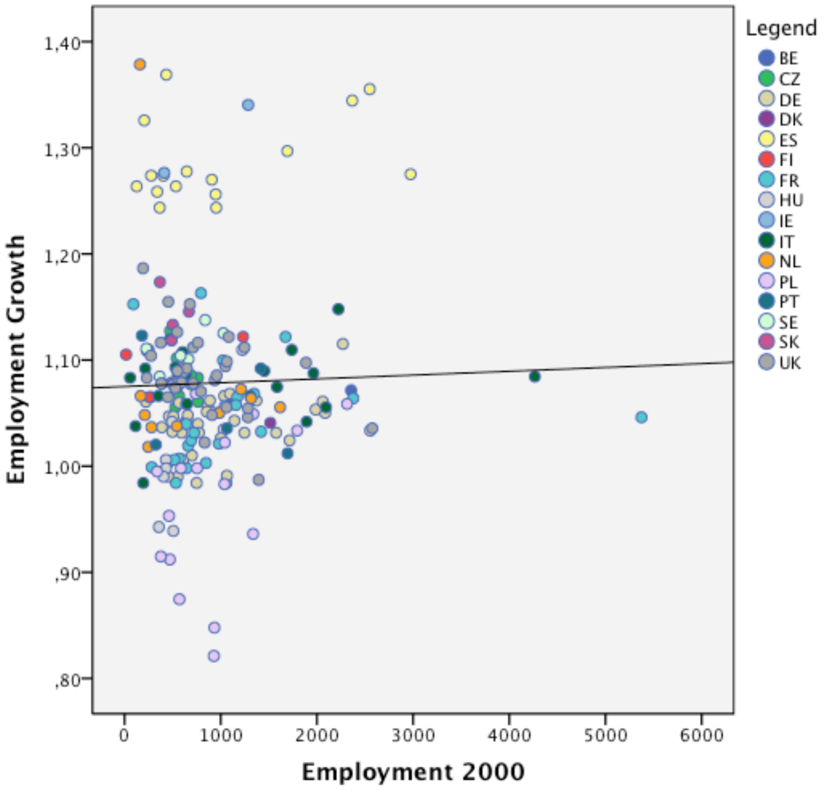
Weber, A. (1929). *Theory of the Location of Industries.* Chicago: The University of Chicago Press.

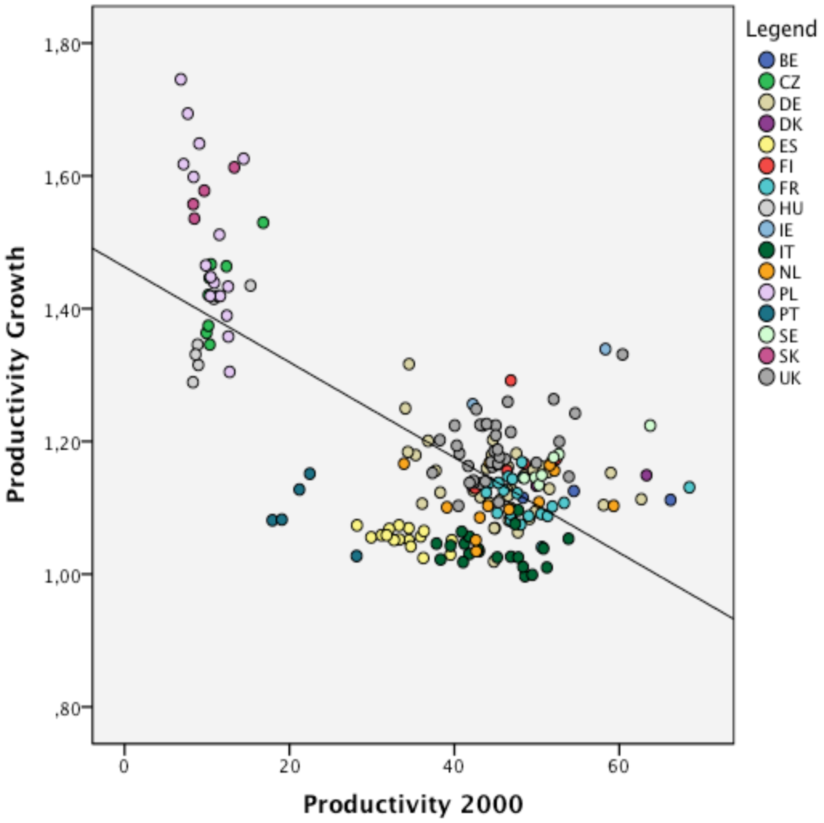
Witte, J. J., (2011). *Sector structure, Evolution and Agglomeration Externalities; Testing the impact of related variety, unrelated variety and specialization in a cross section of European regions.* Utrecht University, Utrecht.

World Bank (2009), *World Development Report: Reshaping economic geography*. Washington, DC: World Bank.

Annexes

Annex 1: Scatter plots of Employment and Productivity, regions specified per country.





Annex 2: Moran's I test for spatial autocorrelation (randomization assumption)

VARIABLE	WEIGHT	I	MEAN	ST,DEV,	Z-VALUE	PROB
UVAR	EU205_2	0,119	-0,005	0,026	4,723	0
PRIV	EU205_2	0,263	-0,005	0,026	10,211	0
PUB	EU205_2	0,015	-0,005	0,026	0,77	0,442
OPEN	EU205_2	0,406	-0,005	0,026	15,705	0
POT	EU205_2	0,638	-0,005	0,026	24,439	0
EDU	EU205_2	0,421	-0,005	0,026	16,199	0
DENS	EU205_2	0,27	-0,005	0,026	10,508	0
WAGE	EU205_2	0,526	-0,005	0,026	20,201	0
OUT00	EU205_2	0,266	-0,005	0,026	10,322	0
EMP00	EU205_2	0,009	-0,005	0,026	0,547	0,585
PROD00	EU205_2	0,658	-0,005	0,026	25,292	0
SPEC	EU205_2	0,389	-0,005	0,026	15,118	0
RVAR	EU205_2	0,053	-0,005	0,026	2,212	0,027
UNEMPGR	EU205_2	0,477	-0,005	0,026	18,379	0
PRODGR	EU205_2	0,627	-0,005	0,026	24,128	0
EMPGR	EU205_2	0,351	-0,005	0,026	13,63	0