DYNREG

Dynamic Regions in a Knowledge-
Driven Global Economy
Lessons and Policy Implications for the EU

WORKING PAPERS

Analysis of Educational Distribution in Europe: Educational Attainment and Inequality Within Regions

Andrés Rodríguez-Pose
Vassilis Tselios

8 / 2007

Analysis of the European educational distribution: educational attainment and inequality

by

Andrés Rodríguez-Pose and Vassilis Tselios

Abstract

The aim of this paper is to visualise and describe the educational attainment and inequality distributions and to detect patterns of global and local spatial autocorrelation, using the European Community Household Panel dataset for 102 regions over the period 1995-2000. It investigates the space-time dynamics of the European educational distributions measured as education level completed and age when the highest education level was completed. This paper also highlights the importance of spatial interaction and geographical location in the human capital performance of the European regions. Without imposing any prior restrictive assumptions on distributions, the exploratory analysis shows that education is geographically autocorrelated due to knowledge and skill diffusion and to the guidelines for education systems and structures which are, as a general rule, set nationally. Thus not only geographical factors such as location, but also institutional ones matter for spatial dependence. The exploratory analysis of the European educational distribution also illustrates the systematic differences between urban and rural areas and between North and South regions. Economies within a cluster interact more with each other than with those outside. Educational attainment is higher in the North and in urban areas, while educational inequality is lower in these areas. Hence spatial dependence and spatial heterogeneity are indeed required features of the European educational analysis.

Keywords: educational attainment, educational inequality, Exploratory Spatial Data Analysis, regions, Europe, urbanisation, EU North-South divide.

1. Introduction

This paper explores and analyses the European educational distribution in terms of educational attainment and inequality, putting emphasis on the role of spatial effects. It aims at investigating the space-time dynamics of the distributions of the average and inequality in education within regions in order to show that spatial autocorrelation and spatial heterogeneity are indeed required features of the European educational analysis. Human capital is expected to be geographically autocorrelated due to some processes, which connect different regions, like educational externalities, and due to institutional factors. This paper emphasises the magnitude of geographical spillover effects in labour market, and highlights the underlying human capital diffusion process, and stress the significance of institutions.

Knowledge diffusion, for instance, is particularly important on regional disparities, because it directly affects regional interactions.

This paper uses data from the European Community Household Panel survey (ECHP) for 102 regions (Appendix A.1) over 1995-2000 so as to examine the way human capital is spatially distributed in the EU and the way in which spatial patterns have probably changed throughout the whole period of study.

From a methodological point of view, this paper is focused on Exploratory Spatial Data Analysis (ESDA) which is a set of techniques aimed at visualising and describing spatial distributions (Baumont, Ertur et al. 2003), such as the distribution of educational inequality. This analysis does not impose any prior restrictive assumption on distributions, while it is applied to ECHP data in order to generate hypotheses about the underlying dynamics of regional economies. It detects patterns of global and local spatial association and suggests spatial regimes and forms of spatial heterogeneity (Haining 1995; Unwin and Unwin 1998; Baumont, Ertur et al. 2003). ESDA highlights the importance of spatial interaction and geographical location in the human capital performance of the European regions. However, it is based on the assumption that the value in the region is spread uniformly throughout the region which is known as 'ecological fallacy' (Cressie 1993).

The first methodological technique of the paper is to map human capital data in order to get a visual view of them and to identify clusters of similar or dissimilar values. Following Jenk's classification, data are divided into six categories (method of natural breaks). The second technique is the application of boxplots which is a commonly used, but very useful method (Fotheringham, Brunsdon et al. 2000). Boxplots will show us the shape of the educational distribution, its central value, and variability. The third technique is spatial autocorrelation analysis which includes tests and visualisation of both global (test for clustering) and local (test for clusters) statistics (Anselin, Florax et al. 2004). It reveals relationships in regional data that could be invisible such as an EU North-South educational divide. Spatial autocorrelation analysis consists of three basic methodological steps. The first step, and the most crucial, is the construction of spatial weights matrices which contain information on the 'neighbourhood' structure for each region. Each region is connected to a set of neighbouring regions by means of a spatial pattern introduced exogenously as spatial weights in order to avoid the identification problems raised by Manski (1993). In this paper, three different spatial weights are considered in order to check for the robustness of the results. However, the appropriate choice of the spatial weights is one of the most difficult and controversial issues in ESDA analysis (Anselin 1988; Florax and Rey 1995; Anselin and Bera 1998; Ertur and Le Gallo 2003). Additionally, the drawbacks of a specific spatial weights matrix are likely to depict the advantages of another one. The specific geographical configuration of the European regions and the choice of the scale of analysis (NUTS I or II) will indeed have some consequences in the choice of the weights matrix (Ertur and Le Gallo 2003). Three different spatial weights schemes are considered. (1) The *rook first order contiguity* matrix. It is constructed in order to reduce the unbalanced connectedness structure of the European regions. (2) The *3-nearest neighbours* matrix. The main advantage of this matrix is that it connects a number of 'islands' such as Sicilia and Sardegna to continental Europe. The Southern United Kingdom is connected to France and parts of Greece to Italy. However, European regions are not very closely connected and compact. (3) The *threshold distance* matrix. The minimum distance required to assume that each region has at least one neighbour is relatively long, because Açores and Madeira are very far from continental Europe. Nevertheless, an advantage of these spatial weights is that there are not unconnected regions. The distance band spatial weights lead to a very unbalanced connectedness structure especially when the spatial units have very different areas, such as the European regions at different NUTS levels. This is because smaller regions have many neighbours, while the larger ones may have very few or none, yielding unconnected observations or 'islands' (Anselin 2003a, 2003b). A major problem on construction of any critical cut-off spatial weights occurs when many regions are missing, since every region must be connected to very other via the spatial weights matrix[1]. For instance, increasing the number of nearest neighbours implies that more regions are affected by the missing observations of the nearest neighbours. Additionally, the 'modifiable areal unit problem' (Openshaw 1983; Arbia 1989; Amrhein 1995) is the basis for the introduction of any spatial weights matrix, because a specific level (NUTS) of spatial aggregation has to be chosen as well as a spatial arrangement in terms of patterns of contiguity or distance (Florax and Rey 1995). The second step is the global spatial autocorrelation analysis. It is not always obvious whether a human capital variable is unevenly distributed over space just by looking at a map. In order to know how strong the spatial association is between neighbouring places, we test, in a statistical sense, for unevenness in the spatial distribution using the most well-known index which is Moran's contiguity ratio or simply Moran's I (Moran 1950)[2]. The third step is the local spatial autocorrelation analysis. Following Fotheringham, Brunsdon et al. (2000), our focus of attention in local analysis is on testing for the presence of differences across regions rather than on assuming that such differences do not exist. These differences exist for many reasons such as the random sampling variations and the misspecification of reality. We use the local variant of Moran's I (Anselin 1995) which is known as Local Indicator of Spatial Association (LISA). This index allows us to depict spatial outliers defined as zones having very different values of an attribute from their neighbours (Fotheringham, Brunsdon et al. 2002: 99). It indicates spatial clustering of similar values

[1] Although the method of 'interpolation' could make predictions for the missing values (Stein 1999), it is not suggested because of the missing national data.

[2] Inference for Moran's I statistic is based on the permutation approach. This is carried out by permuting 999 times the observed values over all locations and by recomputing Moran's I for each new sample.

around the observations (Anselin 1995). The results are illustrated in a cluster map. The cluster map which is a special choropleth map shows those European regions with a significant local Moran statistic classified by type of spatial correlation is generated[3]. Then, the Pearson correlation coefficient is used as a measure of linear association between human capital variables[4]. Finally, ESDA suggests forms of spatial heterogeneity. It allows us to investigate underpinning factors behind educational distribution. We establish links between clusters and underpinning factors. A cluster (or spatial club) is a group of regional economies that interact more with each other than with those outside (Fischer and Stirbock 2006). This method is very useful because economic theory provides no information on the number of regimes or on the way in which foundation factors determine the different clusters of agglomeration (Durlauf and Johnson 1995). This paper examines differences between highly agglomerated urban regions and rural (and usually peripheral) regions; and between the Southern and the Northern regions of Europe.

This chapter is organised as follows. In the next section, definitions of educational attainment and inequalities are presented. Two proxies of educational distribution are used: education level completed, and age when the highest education level was completed. In Sections 3 and 4, ESDA on educational attainment and inequality are displayed, respectively. In Section 5, the relationship between the average and inequality in educational distributions is illustrated. Final section concludes.

2. Defining educational attainment and inequality

A first issue is how to define, to measure and to compare skills, knowledge and competences over time and among regions. This paper explores the formal definition and measurement of two proxies for educational attainment: the average education level completed, and the average age when the highest education level was completed.

Educational attainment can be defined in terms of various human attributes, such as the knowledge, skills and competences embodied in individuals that are relevant to economic activity (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development

[3] The high-high (high surrounded by high) and low-low (low surrounded by low) regions suggest clustering of similar values (positive spatial autocorrelation), whereas the high-low and low-high locations indicate clustering of dissimilar values (negative spatial autocorrelation). The cluster itself consists of the core as well as the neighbours. Anselin (2003a; 2003b) strongly recommends a sensitivity analysis before interpreting the results of LISA maps. More specifically, a 999 permutation procedure at the 0.05 significance level (p-value) is chosen in order to provide stability of the results (Anselin 1995). The tighter significance criterion eliminates some regions from the map.

[4] In the correlation analysis, there is no distinction between the dependent and explanatory variables, while both variables are assumed to be random (Gujarati 2003).

1998). Broadly speaking, measurements of educational attainment could be classified into two basic categories.

The first category describes educational attainment of the population within a society in terms of the percentage who have successfully completed various levels of formal education as defined by the International Standard Classification of Education (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development 1998). The term ‘level’ is defined in relation to the years of study and age associated with an educational cycle. These indicators show how many people have completed each level of initial education. A related measure is the average number of years of schooling completed. It assumes that a year of education will add a constant quantity to human capital stock, whether undertaken by a primary school child or a post-graduate student. Recent studies of measuring human capital stock as the percentage who have gained upper-secondary and tertiary level qualifications or the estimated average number of years spent in completed episodes of primary, secondary and tertiary education are the work of Ram (1990), Barro (1991), Benhabib and Spiegel (1994), Gemmell (1996), Pritchett (1996), Temple (1999) and Ciccone (2004), among others.

The second category provides a relatively novel approach to the measurement of skills and competences consistent with International Adult Literacy Survey. In this respect of human capital stock, adults are tested on three literacy scales (prose, document and quantitative) and assigned to one of five levels of literacy on each scale. The levels represent the varying degrees of complexity in the components of literacy skills needed in different situations. Literacy scores reflect the degree to which adults develop or lose skills initially acquired at school. Fewer studies have put emphasis on the measurement of the quality of educational attainment (i.e. scores on international comparable examinations, talent to engineering, percentage performing at each of five levels of measured literacy in three domains) such as the work of Murphy, Shleifer et al. (1991), Tallman and Wang (1994), Hanushek and Kimko (2000) and Barro (2001).

This analysis is focused on the educational attainment of individuals as measurement of human capital stock, rather than the more complex relationships which combine both the quantity and the quality of human capital endowments within regions. Besides, measurement of human capital stock have been strongly guided by what is possible to measure, rather than by what it is desirable to measure (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development 1998: 89). In this study, two proxies for educational achievement are presented which are aggregate indicators of formal education based on ECHP survey. This, however, implies ‘aggregation biases’ of various sorts and the imposition of restrictions such as homogeneity within regions (Sianesi and Van Reenen 2003). Consequently, some variations in human capital are likely to be lost. Inference about an individual is made using aggregate data for a region.

The first proxy for educational attainment is the average (of highest) education level completed. It considers three grades: less than second stage of secondary education level, second stage of secondary education level, and recognised third education level. Individuals are classified into any of three educational categories which are mutually exclusive. This proxy is collected by the regionalised microeconomic variable '*Highest level of general or higher education completed*', which is extracted from the ECHP dataset. The three levels of the formal education are defined by the International Standard Classification of Education and allow international comparisons. This proxy is based upon two crucial assumptions. The first assumption is that an increase in education level completed adds a constant quantity to human capital stock, undertaken either by a primary or by a secondary student. The second one is that acquisition of postgraduate degrees will not add any quantity to human capital stock, because both graduate and postgraduate degrees belong to the same category ('recognised third level education'). This proxy has been defined by Psacharopoulos and Arriagada (1986) and Ram (1990). The average education level completed is given by the following index:

$$EMN = \sum_j L_j S_j \ ,$$

where $j \in \{1,2,3\}$ are the educational categories, L_j is the proportion of the respondents who fall in the j^{th} category and S_j, at the risk of some oversimplifications, denotes an assessment of each category. More specifically, $S_1 = 2$ for recognised education third level completed, $S_2 = 1$ for second stage of secondary education level completed, and $S_3 = 0$ for less than second stage of secondary education level completed[5].

This proxy, in practice, cannot be compared across European countries with different requirements for completing any given formal educational level. When comparing educational attainment across countries, there is no consistent definition of what a particular level means in terms of knowledge, competences and skills (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development 1998). The completion of a given level can be associated with somewhat different lengths of study in different regions[6]. The duration of some upper secondary and tertiary programmes differs. For instance, there are many short programmes at upper secondary

---

[5] Although the availability of educational categories is very limited (three categories only) and the concept of 'education level' is broad due to differences in national education systems, this assessment is likely to correspond to the numbers of years of schooling, because if the first stage of secondary education level is a base year, the number of years of the second stage of secondary education level is, for most European countries, half than the number of years for recognised third education level. In other words, the minimum duration of getting a second stage degree of secondary education lasts 3 years, as the number of getting a postgraduate University degree (Bachelor degree).

[6] However, the Bologna protocol will reduce the problem of comparability in the future.

level in France (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development 1998). The education systems and structures of each country vary in terms of resources, duration and the preparation of entering students (Sianesi and Van Reenen 2003). For example, the requirements of some knowledge and skill demonstrations in order to pass courses and grades vary widely among countries. Thus, national data on educational attainment are hardly comparable, due to the significant differences in education systems, structures and traditions (Rodríguez-Pose and Vilalta-Bufi 2005). This proxy measures the amount of education undertaken and certifies, in the different context of each European country's education system, acquisition of certain types of knowledge and skill. Additionally, this proxy ignores learning on courses that do not lead to a recognised qualification such as enterprise-based or on-the-job training programmes. Finally, completing a level of education certifies certain knowledge and skills without looking at the time of completion.

The second proxy for educational attainment is the average age of individuals when the highest grade was completed. It is collected by the microeconomic variable '*Age when the highest level of general or higher education was completed*', which also is extracted from the ECHP dataset. This proxy assumes that a year of education will add a constant quantity to human capital stock, whether undertaken by a secondary or tertiary school student. Hence a year of education is a constant unit regardless of level. Furthermore, when assessing the impact of an additional year of education, it is assumed that one year of, for instance, secondary schooling is equivalent to a year at the same grade in other regions and countries. The second proxy is defined as

$$AMN = \frac{1}{N}\sum_{i}^{N} AGE_i \ ,$$

where $i \in \{1,2,...,N\}$ are individuals and AGE_i is the age of the i^{th} individual when the highest education level was completed. This proxy is likely to correspond to differences in duration of studies, but only when there is no any formal educational inactivity period, such as leave. A potential drawback is that it may add short term unemployment and economic inactivity periods to human capital endowments. This proxy also is likely to add training period to human capital stock, only when it has been completed before the highest education level was reached. Hence, it is likely to consider 'wider' definition of human capital investment encompassing experience, learning-by-doing and on-the-job training. Human capital stock as average age is possible to develop indirect measures of the value placed on skills in the workplace and of the benefits to individuals of work-related training. The main point is that age at highest qualification includes any activity prior to final qualification, some of which may be building human capital and some not.

The ideal measures of human capital would be in terms of the output of education, but due to the difficulties of obtaining such measures, input measures tend to be used (Sianesi and Van Reenen 2003: 168). The presented proxies for educational stock are in terms of the input of formal education without considering the output of knowledge, skills and competences embodied in individuals and probable without a wider definition of human capital investment encompassing experience and learning-by-doing (Sianesi and Van Reenen 2003). Completion of educational levels is only broadly associated with some forms of economically-relevant knowledge, skills and competence and does not look at human capital stock attributed directly. A certificate of tertiary education, for example, registers the fact that a student has passed certain courses and exams, but does not certify that he or she has spent a certain amount of time studying (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development 1998: 82). Hence regional differences in the educational attainment cannot explain differences in adult literacy performance. In other words, they do not measure how much in practice attributes are worth in economic terms.

Both proxies do not take into account that skills are lost through disuse. They ignore the depreciation of human capital. Depreciation of skills is often associated with unemployment and economic inactivity. A person's qualifications are kept for life, while the qualities required to gain them may depreciate over time (Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development 1998: 82). The study of Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development (1998) shows that, firstly, in some countries many less-educated people have high literacy, while in others many better-educated ones have low literacy; and secondly, the same level of education yields on average very different literacy outcomes. According to this study, direct skill measures provide a more accurate measure of human capital stock, because better reflects learning, training and skill attrition throughout life. Nevertheless, measuring adult skills directly gives only a partial picture of attributes relevant to economic activity, whereas they do not take into account the depreciation of skills during adulthood.

The 'relative' measures of educational inequality have been used in many studies before (Marin and Psacharopoulos 1976; Winegarden 1979; Ram 1990). In the work of Ram (1990), for example, educational inequality is represented by the standard deviation of the educational distribution for each observation. However, more recently studies use 'relative' measures of educational inequality (i.e. Cornia, Addison et al. 2001; Thomas, Wang et al. 2001; Castello and Domenech 2002; De Gregorio and Lee 2002). Castello and Domenech (2002), for instance, taking school attainment levels, have computed the Gini coefficient. Thomas, Wang et al. (2001) measure inequalities in educational attainment using the education Gini and Theil indices.

In this study, educational inequality is measured using the formula of income inequality indices: the relative mean deviation index, the Gini coefficient, the generalised entropy index and the Atkinson index. Similarly to educational attainment, two proxies for educational inequality are presented.

The first proxy is inequality in education level completed. It is collected by the same variable that the average education level completed has already been measured ('*Highest level of general or higher education completed*')[7]. Theil index takes its minimum value (0) when the entire population is concentrated in a single educational category, while it takes its maximum one ($\log N$) when the entire population belongs to the less than second stage of secondary education level completed category (S_3), except for one person only who has a recognised third level education degree.

The second proxy is inequality in age when the highest education level was completed and is collected by the same variable that the average age when the highest grade was completed has been measured ('*Age when the highest level of general or higher education was completed*')[8]. Educational inequality is

---

[7] Consider a population of individuals $i \in \{1,2,...,N\}$, where each person is associated with a unique value of the measured formal education level completed. It has been assumed that

$$y = \begin{cases} 0 \ \textit{for less than second stage of secondary education level completed} \\ 1 \ \textit{for second stage of secondary education level completed} \\ 2 \ \textit{for recognised third education level completed} \end{cases}$$

such that $\sum_{i=1}^{N} y_i \equiv Y$ [7]. We define the education level completed ratio r_i as the ratio of y_i to the average $\overline{Y}$ ($\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i = \frac{Y}{N}$) $r_i = y_i / \overline{Y}$. By definition, educational equality exists when any education level completed is equally distributed across all persons (all persons hold the same higher degree). Hence, educational inequality is zero when and only when $r_i = 1.0$ for all i; otherwise, inequality is greater than zero. Conceptualising inequality in education level completed as the average disproportionality across all persons implies that the degree of inequality depends on the average distance of the education level completed ratios r_i from 1.0. Educational inequality is unaffected by proportional increases or decreases. Inequality indexes ($EINEQ$) can be expressed in a common form $EINEQ = \frac{1}{N}\sum_{i=1}^{N} f(r_i)$, where f denotes the disproportionality or distance function which captures the mathematical functions for determining deviations of education level completed ratios from 1.0. For instance, using the formula of income Theil entropy index ($GE1$), inequality in education level completed is defined as $EGE1 = \sum_{i=1}^{N} z_i \log(Nz_i)$, where z_i is human capital share that is individual i's higher education level completed as a proportion of total human capital for the entire regional population.

[8] This index ($AINEQ$) can be expressed in the form $AINEQ = \frac{1}{N}\sum_{i=1}^{N} f(r_i)$, where f denotes the distance function which captures the mathematical functions for determining deviations of ratios of age when the highest education level was completed from 1.0. Using, once again, the formula of income Theil entropy index ($GE1$), inequality in age when the highest education level was completed is defined as $AGE1 = \frac{1}{N}\sum_{i} r_i \log(r_i)$.

zero when and only when all people have completed their highest education level at the same age; otherwise, inequality is greater than zero.

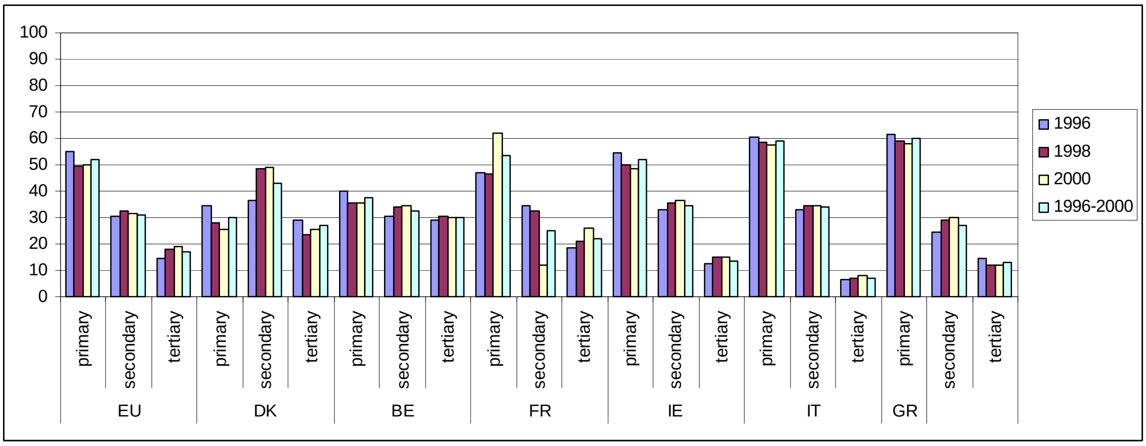
To sum up, the analysed proxies for educational attainment and inequality are more measurements of the quantity and availability of a region's human resources (input measures), rather than measurements of the quality of human capital endowments (output measures). The quality of education is not taken into consideration. Measuring the quantity of education is only a crude measure of skill differences (Hanushek and Kimko 2000).

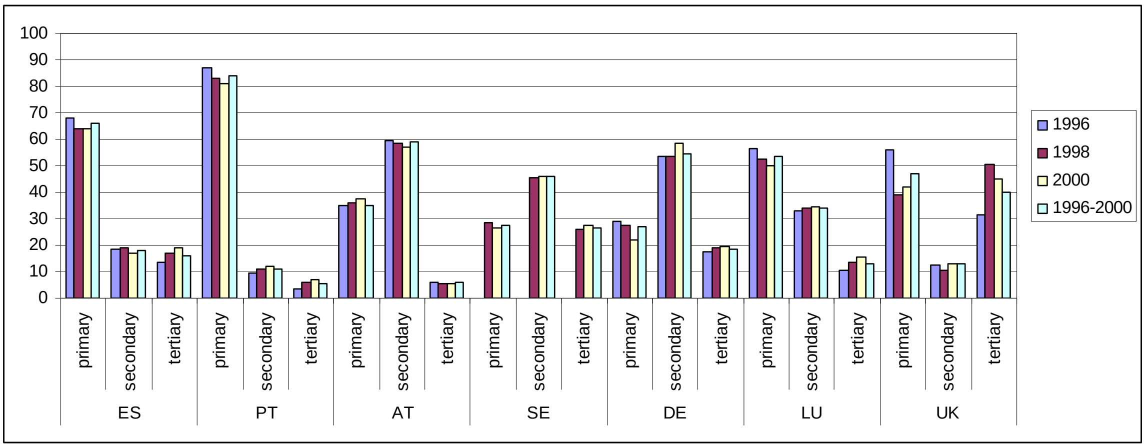
3. Exploratory Spatial Data Analysis on educational attainment (1995-2000)

We firstly consider educational attainment as average education level completed. Regarding human capital stock as a quantitative variable, we investigate ESDA on human capital endowment within European regions. However, a preliminary analysis of national educational attainment is obtained, exploring human capital as a qualitative variable.

According to the International Standard Classification of Education, the educational attainment of the respondents within Europe is explored in terms of the *percentage* of people who have only completed the primary (or the first stage of secondary education), the secondary or the tertiary education level. Between 1995 and 2000, 48% of European respondents who have completed formal education hold a secondary education level degree and 17% of them have also completed tertiary education. Figure 1 demonstrates the recent evolution of educational attainment by country and formal education level completed. The results show that first Portuguese and then Spanish citizens are the least educated in Europe, whereas Denmark, Sweden and Belgium have the highest and also the most equally distributed human capital endowments. Danish, Swedish and Belgian citizens may have, for example, a higher level of aspiration and have put more effort into their career (Hansen 2001). They may have maximised their economic welfare by investigating a higher amount in human capital (Becker and Chiswick 1966). Hansen (2001), however, says that the higher education which has been attained by a high proportion of the Swedish and British population is likely to lead to inflation in the value of educational credentials. According to Figure 1, Italy, Portugal and Austria have the lowest percentage of high-educated people. Ireland and Luxembourg's segmented distribution of educational achievement follows the European distribution. The percentage (47%) of British who have completed only the primary education is high and close to the percentage (42%) of ones who hold a certificate on higher education (tertiary). This is likely to demonstrate polarisation of educational attainment which means an increase in the homogeneity within groups of education levels, but an increase also in the distance between groups. The distance between the primary and the tertiary education level completed is likely to depict the gap between the individual's effort into their career or their lifetime economic opportunities. Between 1995 and 2000, the component of human capital stock at different education levels remained almost the same for secondary education and increased slightly for higher education (14.6% in 1996, 17.9% in 1998 and 19.8% in 2000). Nevertheless, the cross-country differences on the percentage at education level completed are significant.

Figure 1: Percentage of respondents with primary, secondary or tertiary education level completed by European country in 1996, 1998 and 2000





Assessing each educational level (as has been mentioned, 0 for second stage of secondary education level completed; 1 for second stage of secondary education level completed; and 2 for recognised third education level completed), human capital stock is transformed into a quantitative variable. Calculating the average education level completed of all European citizens, the European educational attainment has somewhat increased. For instance, it increased from 0.5 in 1996 to 0.7 in 2000.

Mapping the average education level completed, it allows us to establish whether educational attainment within regions is randomly distributed over the EU or there are similarities between regions. Figure 2 shows the spatial distribution of the average education level completed within regions in 1996, 1998 and 2000. There are striking disparities in human capital endowments between different regions of Europe. In Portugal, Spain, Italy, and Greece the average education level completed is lower than anywhere in the Union. Educational attainment is approximately half of the EU average in these countries. It is well above average in Northern Europe including the United Kingdom, Denmark, Sweden, Belgium, and Germany. Northern regions with relatively high human capital endowment are and remain localised close to other regions with relatively high human capital endowment; while, Southern regions with relatively low human capital endowment are and remain localised close to other regions with relatively low human capital endowment.

Educational attainment disparities seem to be higher at national level than at subnational one, because the guidelines for education systems and structures are, as a general rule, set nationally (Rodríguez-Pose and Vilalta-Bufi 2005: 552). European regions have to comply with national guidelines and curricula (Rodríguez-Pose and Vilalta-Bufi 2005: 552). Most institutions, even private or religious schools, are under the control of national governments and usually funded by government expenditures. For instance, university fees are generally set nationally. Nevertheless, within the United Kingdom and Germany there are striking regional disparities demonstrating human capital segregation. More specifically, in the United Kingdom, the educational attainment measured as average education level completed is highly concentrated in Southern (Bedfordshire, Hertfordshire, Berkshire, Buckinghamshire, Oxfordshire, Essex, Hampshire, Isle of Wight and Kent) and Northern (Scotland) regions; and in Germany, human capital endowment is higher in North-Eastern regions (Brandenburg, Mecklenburg-Vorpommern, Berlin, Sachsen, Sachsen-Anhalt and Thüringen). German regions are likely to have signs of powers over a devolved education system as the subnational disparities of educational attainment illustrate. Besides, the German public schools are subject to state laws, not federal, which is why there are considerable differences between states[9]. British and German regional disparities may have to do with the spatial level of analysis, since the aggregation level in the United Kingdom and Germany is NUTS II. However, data that are close together in space (i.e. NUTS II) are more often alike than those that are relatively far apart (i.e. NUTS I) (Cressie 1993). Regions in NUTS I level may be too large and the unobserved heterogeneity may create an ecological fallacy. British and German disparities probably also arise from boundary mismatching between NUTS II and the actual market boundaries over which economic processes operate.

Considering the urbanisation level, human capital endowment is higher in city-regions (Greater London, Île de France, Région de Bruxelles) than anywhere. These cities are likely to attract high-qualified migrants in search of better working prospects. Many people move to core cities in search of better educational opportunities, employment, further career prospects and standards of living. The higher education institutions are generally located in cities. The local provision of higher education

[9] www.watzmann.net

institutions may itself contribute to a growth in the local stock of human capital (Bennett, Glennerster et al. 1995). Educational stock has important effects on the structure of the local economy, either city or region. However, the existence of high-qualified institutions in a region or city is not sufficient to ensure a high human capital endowment. The ability of higher education infrastructure to increase the stock of human capital within a regional market depends on the ability of the region to attract as well as to retain high quality students and workers (McCann and Sheppard 2001). The institutions of the European cities seem to attract students of sufficient learning ability and the urban labour market may retain them once they have graduated. This outcome depends on the previous migration history of the individual (Davanzo 1976) and on the personal unemployment (Davanzo 1978) (McCann and Sheppard 2001: 137). High-educated workers are more likely to make the necessary movements required in order to achieve higher promotion. Additionally, they are prone to migrate more as a way of achieving greater employment returns. These findings are consistent with Fingleton's (2003) findings. According to him, although there are undoubtedly variations due to varying national education systems, structures and traditions, it is however revealing that regions with high levels of educational attainment are those urbanised, non-peripheral regions which one would consider to be the productive core of Europe (Fingleton 2003: 12). He also says that '*regions specialised in high value added manufacturing, research and development and service activities will also have work-forces with commensurate skills*' (Fingleton 2003: 13).

Figure 2: Spatial distribution of average education level completed (EMN) in 1996, 1998 and 2000

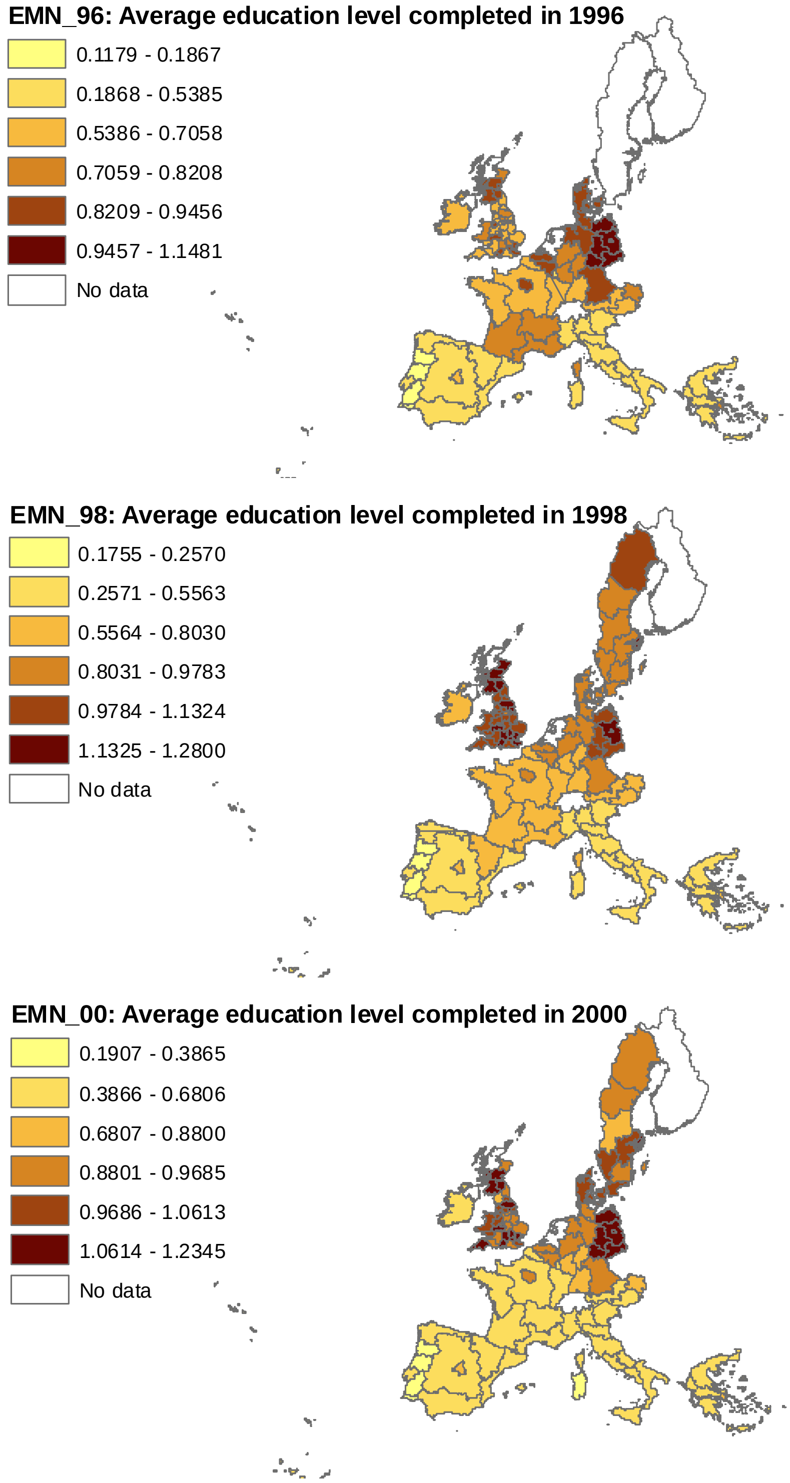
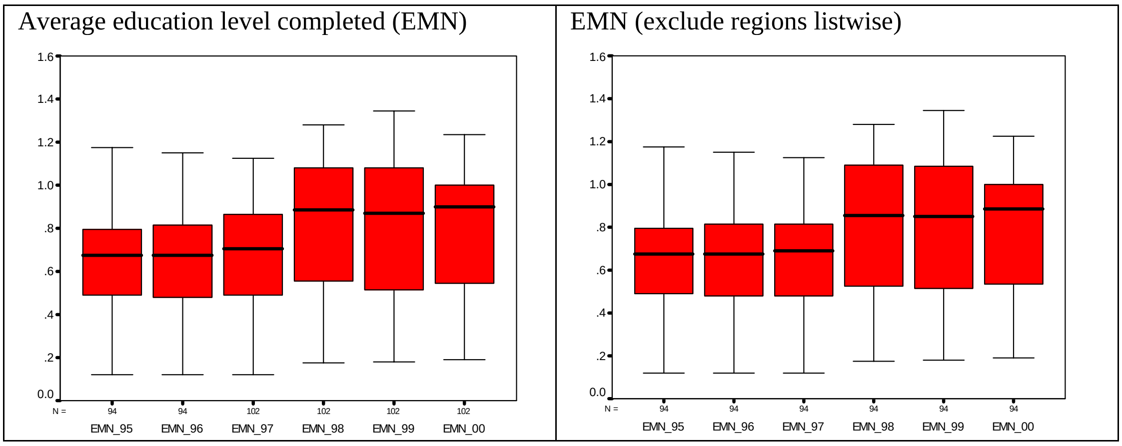


Figure 3 illustrates the boxplot for the average education level completed within European regions from 1995 to 2000. Although the segments of education are unequally distributed over space, there are not outliers. This is a sign of the compactness of the European distribution of educational attainment. The median remained constant between 1995 and 1997, and between 1998 and 2000 (0.89 in 1998, 0.87 in 1999 and 0.90 in 2000), but increased significantly (by 0.19) from 0.70 in 1997 to 0.89 in 1998. The average had the same evolution. Furthermore, the interquartile range increased from 1997 to 1998, indicating increased variability of the average education level completed. The interquartile range and the variations in the whiskers are somewhat longer in 1999. It depicts that human capital endowments cover a larger spectrum. Finally, the European distribution of the average education level completed accepts normality over the period 1995-1999, but rejects it in 2000. The ratio of skewness to standard error is negative which indicates a left tail[10].

Figure 3: Boxplot for average education level completed (EMN)



Short trends in the evolution of human capital disparities across the EU can be captured not only by distribution maps and boxplots, but also by simple statistical measures of spatial dependence, such as Moran's I test statistic. Constructing the rook first order contiguity spatial weights matrix for average education level completed, Moran's I global spatial autocorrelation statistics are high (Table 1). These statistics depict that there is a high positive spatial autocorrelation of human capital endowment. Considering the space-time correlation, it is shown that Moran's I statistic between a region's human capital endowment in 1998 and neighbouring regions' endowment in 1996 (which is the space-time correlation of human capital stock in 1996) is 0.5547 when Sweden is excluded, and the space-time correlation in 1996 is 0.6896. Both space-time correlation statistics depict a positive spatial correlation. Moran's I statistics computed using the 3-nearest neighbours spatial weights matrix are also high. Finally, the threshold distance schemes depict a positive spatial autocorrelation as well, but it is lower than using the other schemes. For instance, the spatial autocorrelation in 1999 is 0.3802 only, including all counties. However, the standardised values of Moran's I statistic appear to be very high possible indicating, once more, a spatial scale problem (Ertur and Le Gallo 2003: 64). Considering the evolution of Moran's I test statistic between 1995 and 2000 shows that the standardised values of the statistic remain approximately the same over the whole period. It indicates a global significant tendency toward

[10] The ratio of skewness to standard error is -0.63 in 1995, -0.91 in 1996, -1.69 in 1997, -1.78 in 1998, -1.20 in 1999 and -2.13 in 2000.

geographical clustering of similar regions in terms of average education level completed. Moran's I statistics lead to the same results for the sign (positive) and significance of global spatial dependence, highlighting the robustness of the results, with regard to the choice of the spatial weights matrix. The statistics show the stock of human capital endowment in a particular region may contribute to output gains in adjoining regions (Lall and Yilmaz 2001). To put it in a slightly different way, statistics are likely to underline the importance of external economies that cross the weak regional boundaries (Vaya, Lopez-Bazo et al. 2004) on the one hand, and the institutional differences between countries which means regions within countries are similar on the other.

Table 1: Moran's I for average education level completed (EMN)

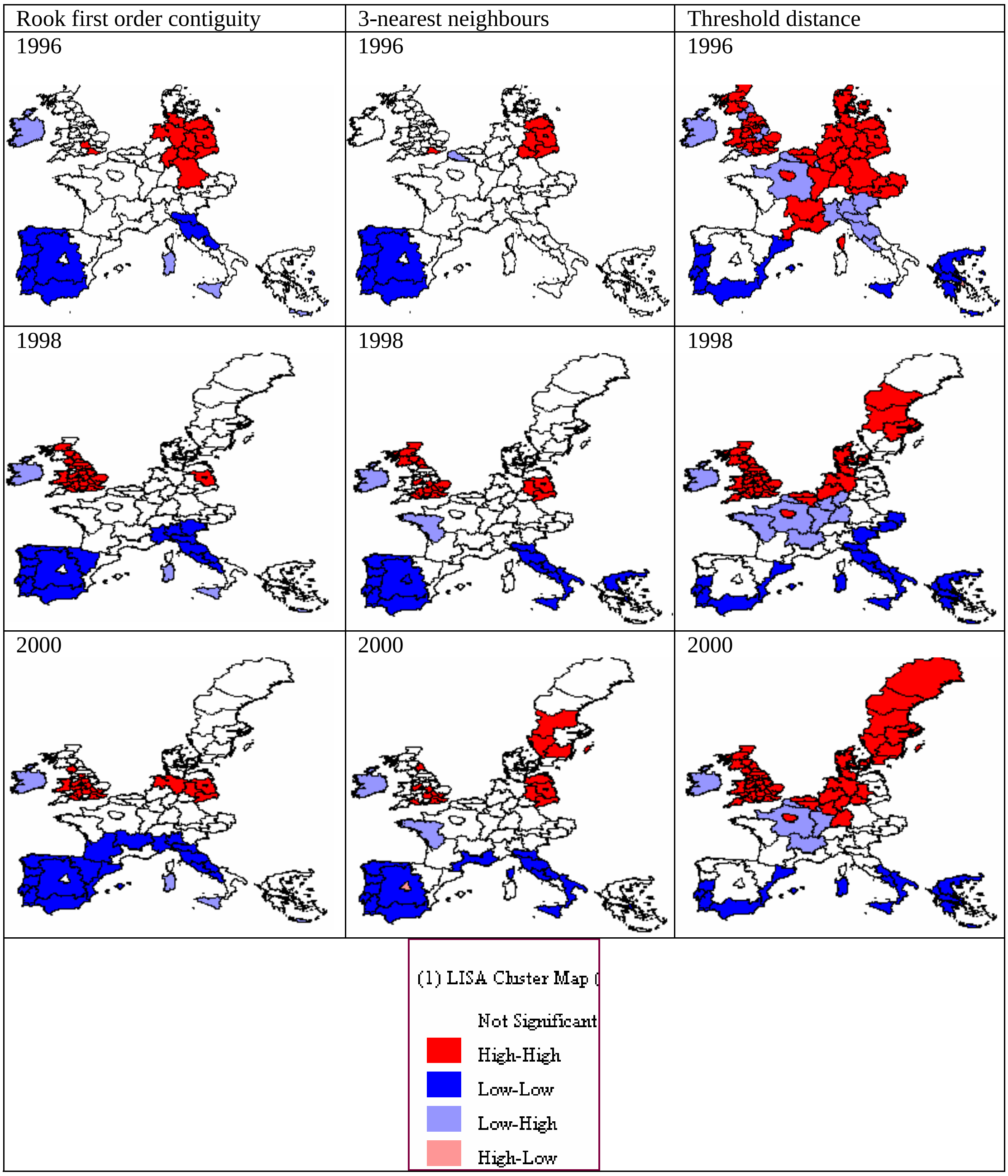
		13 countries (E[I]=-0.0099)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995												
	1996												
	1997	0.6175	-0.0084	0.0776	8.0657	0.7617	-0.0119	0.0754	10.2599	0.4139	-0.0091	0.0225	18.8000
	1998	0.7313	-0.0107	0.0727	10.2063	0.8250	-0.0102	0.0747	11.1807	0.4080	-0.0096	0.0217	19.2442
	1999	0.7503	-0.0088	0.0790	9.6089	0.8002	-0.0118	0.0747	10.8701	0.3802	-0.0088	0.0226	17.2124
	2000	0.6900	-0.0039	0.0746	9.3016	0.7752	-0.0104	0.0751	10.4607	0.3968	-0.0114	0.0215	18.9860
Space-time correlation	1998												
	2000	0.6896	-0.0103	0.0725	9.6538	0.7793	-0.0106	0.0741	10.6599	0.3963	-0.0110	0.0212	19.2123
		Excluded SE (E[I]=-0.0108)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995	0.6109	-0.0082	0.0756	8.1892	0.7466	-0.0093	0.0768	9.8424	0.3491	-0.0115	0.0226	15.9558
	1996	0.6119	-0.0079	0.0746	8.3083	0.7433	-0.0101	0.0751	10.0320	0.3577	-0.0105	0.0225	16.3644
	1997	0.6085	-0.0068	0.0768	8.0117	0.7384	-0.0080	0.0762	9.7953	0.3619	-0.0126	0.0225	16.6444
	1998	0.7419	-0.0081	0.0772	9.7150	0.8297	-0.0072	0.0768	10.8971	0.4061	-0.0110	0.0211	19.7678
	1999	0.7607	-0.0112	0.0773	9.9858	0.8039	-0.0147	0.0749	10.9292	0.3770	-0.0118	0.0219	17.7534
	2000	0.7009	-0.0093	0.0775	9.1639	0.7776	-0.0062	0.0809	9.6885	0.3837	-0.0098	0.0234	16.8162
Space-time correlation	1998	0.5547	-0.0061	0.0676	8.2959	0.6534	-0.0100	0.0709	9.3568	0.3396	-0.0085	0.0219	15.8950
	2000	0.7029	-0.0131	0.0716	10.0000	0.7850	-0.0063	0.0729	10.8546	0.3933	-0.0119	0.0217	18.6728

Note: All statistics are significant at p=0.001.

Moran's I statistic does not allow to assess the regional structure of human capital spatial autocorrelation. LISA are used to test the assumption of a random distribution by comparing the human capital values of each specific region with the values in the neighbouring regions (Ertur and Le Gallo 2003). Figure 4 illustrates the cluster maps for average education level completed in 1996, 1998 and 2000, at three weighting schemes. They show the local variation of educational attainment in spatial autocorrelation. Different trends in human capital distribution exist across regions in the EU. The weighting schemes of the first order contiguity and the 3-nearest neighbours show that clusters of regions with poor human capital endowments are found across Italy, in Southern France (in Sud-Ouest and Centre-Est considering the first order contiguity schemes, and in Méditerranée for the 3-nearest neighbours ones) in 2000, in Portugal and in Spain. Conversely, two clusters of regions with high human capital stock can be found in Southern England and in Eastern Germany (Berlin, Brandenburg and Sachsen-Anhalt). The distance band weights schemes depict more expanded clusters. For instance, the high-level of education cluster of the United Kingdom includes all regions, in 1998 and 2000.

Furthermore, many regions in Central Europe are spatial outliers such as Northern Italy in 1996, and French regions of Bassin Parisien, Nord - Pas-de-Calais, Est and Centre-Est in 1998 and 2000. Finally, this figure confirms the fact that the average education level completed is higher in Southern Europe. A cluster of rich human capital regions (the North) is distinguished from a cluster of poor human capital regions (the South).

Figure 4: Cluster map for average education level completed (EMN) in 1996, 1998 and 2000

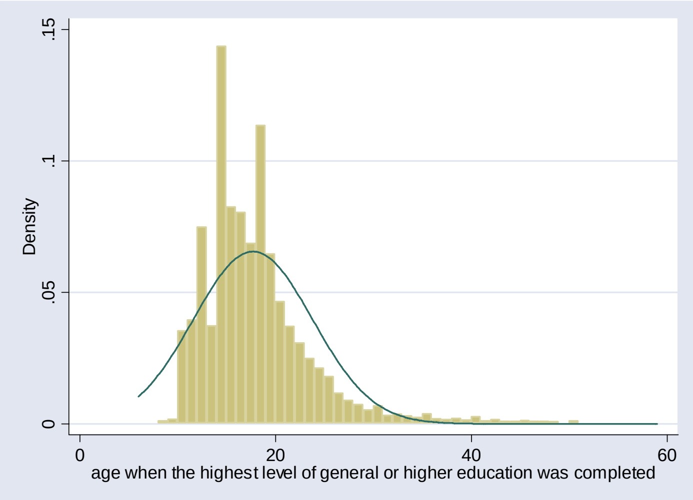


Generally speaking, the results reveal the persistence of human capital disparities among the European regions over time following the urban-rural and the North-South polarisation. This detects two forms of spatial heterogeneity. In other words, the findings show that economic behaviour is not stable over space. The spatial regimes can be linked by several results from regional development theories, such as the New Economic Geography (NEG) models (i.e. Krugman 1991a, 1991b; Krugman 1993; Krugman and Venables 1995, 1996; Puga and Venables 1996; Martin 1998; Fujita, Krugman et al. 1999; Martin 1999a, 1999b; Martin 1999c; Puga 1999; Fujita and Thisse 2002) which stress the significance of spatial effects, via, for instance, home market (Krugman 1980; Helpman and Krugman 1985; Davis and Weinstein 2003) and price index (Fujita, Krugman et al. 1999) effects; and the cumulative causation theories (i.e. Rosenstein-Rodan 1943; Perroux 1950; Myrdal 1957; Hirschman 1958; Kaldor 1970, 1981, 1985; Arthur 1994). If one Northern region acts to attract human capital, all Northern regions benefit from the spillovers. Nevertheless, the spatial clustering is likely to correspond national institutional differences such as in educational system. Therefore, spatial autocorrelation and spatial heterogeneity are unavoidable features of human capital variation analysis.

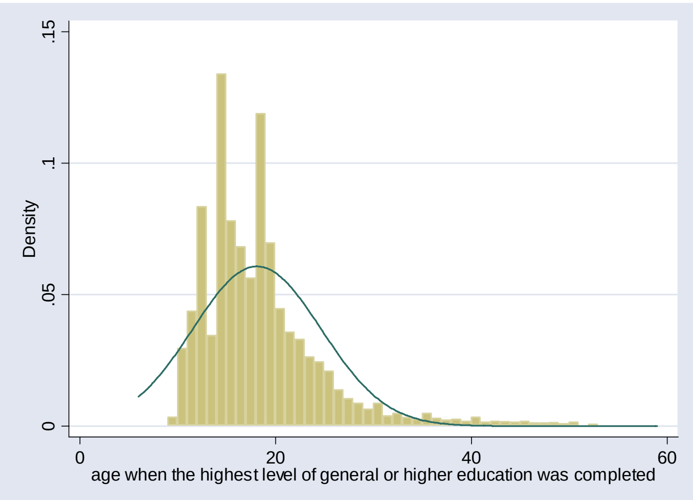
The European distribution of age of respondents when the highest grade was completed is illustrated by the following histograms in 1996, 1998 and 2000 (Figure 5). All histograms appear two peaks; one at age 15 and another at age 19. This generally corresponds with the age of completion of the first and the second stage of secondary education level. Another smaller peak appears at age 13 in which people complete the primary school. After the age of 19, the European age distribution follows the normal distribution. Comparing the histograms, it is shown that the peak at age 15 is lower in 1998 than in 1996, denoting that people continue their studies. Most respondents have completed their highest level of general or higher education when they were between 15 and 20 years old.

Figure 5: Histogram of age of respondents when the highest education level was completed

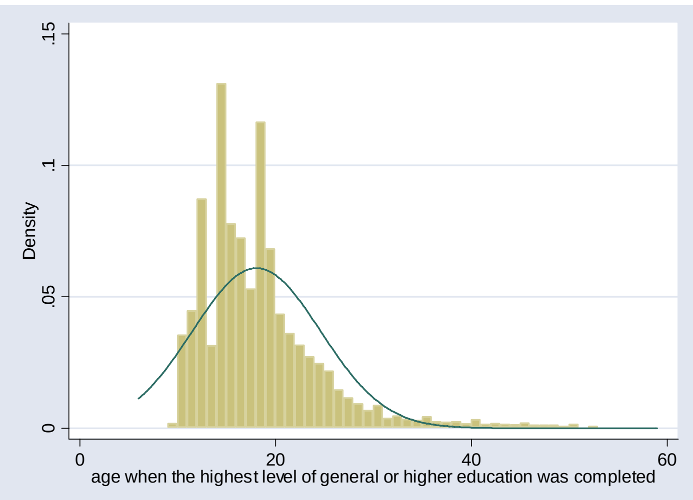
1996



1998



2000



The spatial distribution, the boxplot, the Moran's I statistic, and the cluster maps for average age when the highest education level was completed are displayed in Appendices A2-A5.

Regional patterns of this proxy for educational attainment are similar to the first one. More specifically, although the number of regions included in 1998 and 2000 are not satisfactory enough because there is no data for France, human capital endowment differs among countries and regions. The geographical distribution of the European human capital endowment is highly clustered (Appendix A.2). German and Danish citizens have completed their educational studies at an older age than any other European citizen. Dig a little deeper, in Germany schools, for instance, attendance is compulsory for children of ages 7 to 18. At least nine years of this period, they must attend a full-time school and then they choose either to continue the full-time schooling or attend a vocational school part-time[11]. Taking into account the variable '*Age when full-time education was stopped*'[12], most German regions and some British ones (i.e. Berkshire, Dorset and Greater London, in 2000) have the highest average age when the full-time education was stopped, highlighting the high human capital endowment in these regions. Furthermore, the difference between the average age when the highest grade was completed and the average age when the full time education was stopped is higher in German regions (i.e. Sachsen, Brandenburg, Sachsen-Anhalt and Berlin, in 2000). The findings do not support the fact that the high human capital stock in Germany might be due to the high portion of part-time students. The duration of studies in German institutions is among the highest in Europe. For instance, the nominal duration of studying physics is 5 years[13]. The spatial distribution of the average age seems to be randomly distributed across the United Kingdom regions, while in Italy and Germany it seems to be concentrated in particular areas. In Italy, there is a North-South divide where high human capital endowments are concentrated on the North and, in Germany, human capital is concentrated in Eastern region. Additionally, Portugal and Greece have the lowest average age in Europe. To sum up, Europe is characterised by wide disparities in the average levels of age when the highest grade was completed.

The boxplot for the average age of individuals when the highest grade was completed (Appendix A.3) shows that German regions and Denmark are outliers and extreme cases. More particularly, the educational attainment of Berlin, Brandenburg, Mecklenburg-Vorpommern and Sachsen is double of the EU average. The distributions are skewed and much of the skewness is due to the outliers and extreme regions in the upper end of the distributions. The skewness is higher in 2000 which denotes that people continue their studies at higher education levels. Nevertheless, the median and the box

---

[11] www.watzmann.net

[12] This variable is available from 1998 to 2001.

[13] www.zhr.rwth-aachen.de

length remained the same between 1995 and 1997, and between 1998 and 2000[14]. The average increased slightly from 18.25 in 1995 to 18.81 in 2000. The distribution of this proxy for educational attainment rejects the normality assumption, because the ratio of skewness to its standard error is greater than +2 which indicates a long right tail[15]. Average age when the highest education level was completed has a significant positive spatial autocorrelation and space-time correlation (Appendix A.4). The standardised values of Moran's I statistic remained almost constant over the whole period of study. The choropleth maps of the first order contiguity weights and the 3-nearest neighbours are quite similar (Appendix A.5). Low human capital endowment is concentrated in Greece (mainly in Voreia Ellada) and in Lazio (in 1998), while Germany is characterised by high human capital stock. Considering the 3-nearest neighbours spatial schemes, Noroeste is the 'core' of another cluster of low human capital endowment. The spatial distribution of educational endowment remained almost the same. The distance band schemes depict expanded poor clusters including Portugal, Spain, Western France, Greece and the Western United Kingdom and, also, an expanded rich cluster including Germany and Denmark. Between the two clusters, there is a Low-High cluster stretching from Eastern France to Italy in which low regions are surrounded by high ones. Once again, the results reveal the persistence of human capital disparities among European regions over time following the urban-rural and the North-South polarisation. The variation in human capital endowment is influenced by region specific characteristics and the availability of high-educated labour in neighbouring Southern or Northern regions. However, this pattern is less intensive than considering educational attainment as average education level completed.

The relationship between the average education level completed and the average age when the highest education level was completed is explored by a cross-tabulation analysis, the comparison of their boxplots (standardised distributions), the Pearson correlation and the bivariate measures of spatial association.

First, the relationship between the age of respondents when the highest education level was attained and the three levels of formal education is analysed by a cross-tabulation analysis. A categorical variable with 6 educational categories (age bands) is created. It has been mentioned that the completion of a given educational level can be associated with somewhat different lengths of study in different countries and thus different age bands. Additionally, comparing educational attainment across countries, there is no consistent definition of not only what a particular level means in terms of

---

[14] The median is 17.54 in 1995, 17.51 in 1996, 17.53 in 1997, 17.87 in 1998, 17.85 in 1999 and 17.73 in 2000.

[15] The ratio of skewness to standard error is 5.49 in 1995, 5.43 in 1996, 5.54 in 1997, 5.36 in 1998, 5.12 in 1999 and 5.44 in 2000.

knowledge and skills, but also what a particular age band means in terms of education level completed. The duration of educational (i.e. tertiary) programmes by educational category (i.e. type of degree) differs between countries. For instance, the minimum period of registration for Bachelor students in Economics is three years full-time in the United Kingdom, while it is four years full-time in Greece. Additionally, the duration, for example, of tertiary programmes differs within countries. In Greece, the minimum period of registration for undergraduate students fluctuates from four (i.e. studying Economics) to six years (i.e. studying Medicine). Educational categories possibly eliminate the requirements of some knowledge and skill demonstrations in order to pass courses and grades. Educational categories also do not distinguish students by full-time or part-time registration. Therefore, in order to check the sensitivity of the results, a second categorical variable (age band) is created which is lagged by one year of the first categorical variable. Generally speaking, in the first categorical variable, the educational categories denote:

- less than 13 (or less than 12): no education level completed;
- 13-15 (or 12-14): primary education completed;
- 16-18 (or 15-17): less than second stage of secondary education level completed;
- 19-22 (or 18-21): second stage of secondary education level completed;
- 23-30 (or 22-29): tertiary education level completed;
- more than 30 (or more than 29): other education level completed.

Table 2 shows that the higher the age of respondents, the higher the education level completed. Considering the first age band, 45.95%, 45.03% and 44.15% of respondents who have completed a less than second stage of secondary education level in 1996, 1998 and 2000, respectively, completed their formal studies when they were between 13 and 15 years old. Taking into account the second age band, 45.57%, 47.02% and 47.10% of respondents who also have completed a less than second stage of secondary education level in 1996, 1998 and 2000, respectively, completed their studies when they were between 12 and 14 years old. The highest portion of respondents who have completed a second stage of secondary education belongs to the age band 16-18 (i.e. 50.10% in 1996) or 18-21 (i.e. 63.54% in 1996). Finally, according to the first age band, 43.32%, 45.08% and 45.70% of European citizens who have acquired a recognised third education level in 1996, 1998 and 2000, respectively, completed their formal studies when they were between 23-30 years old. Considering the second age band, the highest portion is between 22-29 years old (i.e. 52.86% in 1996).

Table 2: Percentage of respondents by age bands and levels of formal education in 1996, 1998 and 2000

	1996			1998			2000		
	less than second stage of secondary education level completed	second stage of secondary education level completed	recognised third education level completed	less than second stage of secondary education level completed	second stage of secondary education level completed	recognised third education level completed	less than second stage of secondary education level completed	second stage of secondary education level completed	recognised third education level completed
<13	27.39	0.17	0.03	30.05	0.07	0.01	32.06	0.13	0.01
13-15	45.95	2.02	1.51	45.03	0.93	2.15	44.15	1.19	2.38
16-18	19.31	50.10	9.91	17.46	45.22	11.44	17.49	45.02	11.74
19-22	3.22	35.46	37.32	3.34	37.75	31.38	2.61	38.21	29.77
23-30	1.84	7.51	43.32	1.70	9.74	45.08	1.64	9.56	45.79
30>	2.28	4.75	7.90	2.42	6.30	9.92	2.04	5.89	10.31
<12	13.95	0.08	0.02	14.40	0.02	0.01	15.56	0.04	0.01
12-14.	45.57	0.75	0.40	47.02	0.23	0.44	47.10	0.36	0.40
15-17	29.75	20.16	6.24	27.88	13.42	8.52	28.21	12.77	9.59
18-21	6.19	63.54	30.57	6.11	66.62	25.11	5.07	67.98	23.50
22-29	2.00	10.00	52.86	1.89	12.42	53.66	1.73	11.99	53.90
29>	2.55	5.47	9.91	2.69	7.30	12.27	2.32	6.86	12.60

Second, considering the boxplots of proxies for educational attainment (Appendix A.6), the median gap between the two proxies becomes even higher from one time period to the next. This probably depicts the decreasing correlation between the two proxies through time for regions which are close to the European average. Additionally, the average education level completed distributions are more skewed than the average age when the highest education level was attained due to the outliers and extreme values. In 2000, the distribution of the average age is skewed on the left. Third, measuring the Pearson index, a positive linear correlation is shown (Appendix A.7). This correlation is higher between 1995 and 1997 than in 1997 and 1998. Fourth, the correlation between the average education level completed within a region and the average age when the highest education level was completed in neighbouring regions, and vice versa, are explored. In 1996, for instance, the bivariate Moran's I statistic between the average education level completed within a region and the average age of neighbouring regions is 0.4534, while between the average age within a region and the average education level completed of neighbouring regions is 0.4918, for the first order contiguity spatial weights schemes, 0.5428 and 0.5457, respectively, for the 3-nearest neighbours weights schemes, and 0.2129 and 0.2140, respectively, for the threshold distance band weights schemes. No matter what the proxy for educational attainment is, geographical location is important in accounting for the human capital performances of the regions due to spatial interactions between regions. The spatial distribution of education stock seems to be far from random.

4. Exploratory Spatial Data Analysis on educational inequality (1995-2000)

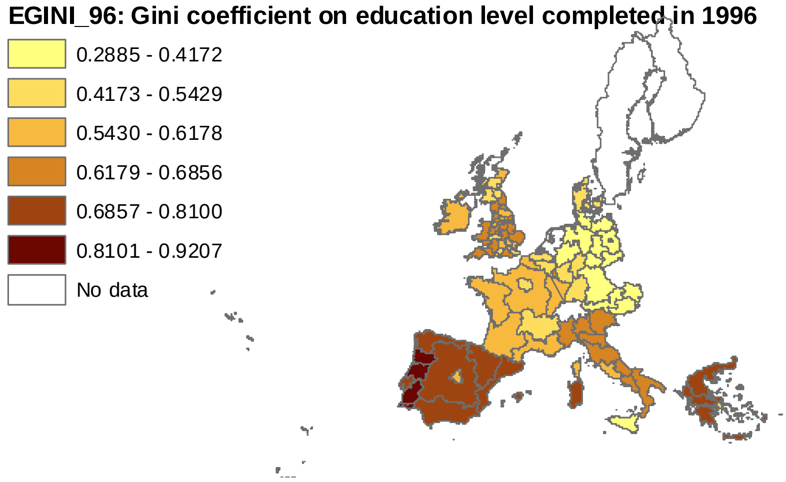
Inequality in education level completed is measured by the Gini index ($EGINI$), the relative mean deviation index ($ERMD$), the generalised entropy index for two different parameters ($EGE1$ when

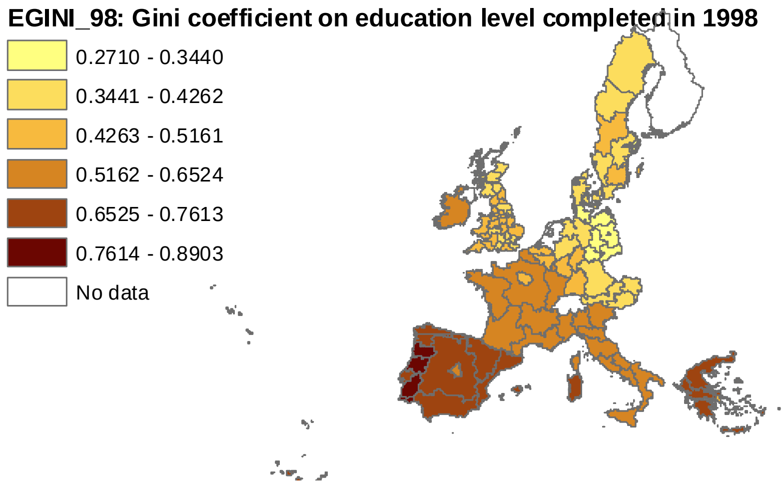
$a = 1$, and $EGE2$ when $a = 2$) and the Atkinson index for one parameter only ($EA050$ when $\varepsilon = 0.50$).

Considering the geographical distribution of the Gini coefficient on education level completed in 1996, 1998 and 2000 (Figure 6), there are striking differences in educational inequality within regions between different parts of the EU. Inequality in human capital endowments is higher in Southern Europe, expanding from Greece, to Italy, Spain and Portugal than in the Northern periphery. The within-region human capital inequality is typically half of the EU average in Germany, Denmark and Sweden. The EU North-South divide shows that regional economies within a Southern group seem to interact more with each other than those outside. Short trends in the evolution of inequality in education level completed demonstrate that inequality remained almost constant, except France and Italy which increased even more in 1998 and 2000. Considering the urbanisation level of each region, educational inequality is lower in the Northern metropolitan areas such as London, Paris, Hamburg and Bruxelles, as well as in the Southern ones such as Madrid, Lisbon and Athens. Additionally, inequality is lower in the metropolis than in the remainder of the respective countries. High-educated workers of rural areas are likely to move to core cities in order to achieve promotion and greater employment returns. Urban market seems to have the ability to attract and retain high quality students and workers. Better educated people move to large cities in search of employment and higher standards of living. Higher human capital individuals tend to be more migratory. The Northern metropolitan areas are getting the most-educated segment of the EU population. Urbanisation, consequently, seems to generate new requirements for the development of higher education. To sum up, the EU North-South divide and the urbanisation degree seem to affect educational inequality. The geographic distributions of other measures of inequalities such as the relative mean deviation index, the Theil index, the squared coefficient of variation and the Atkinson index yield similar results[16].

[16] The results are provided under request.

Figure 6: Spatial distribution of Gini coefficient on education level completed (EGINI) in 1996, 1998 and 2000





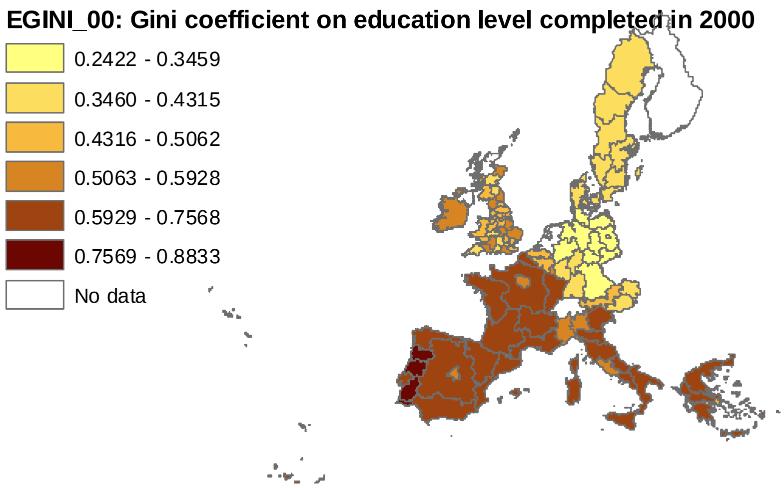
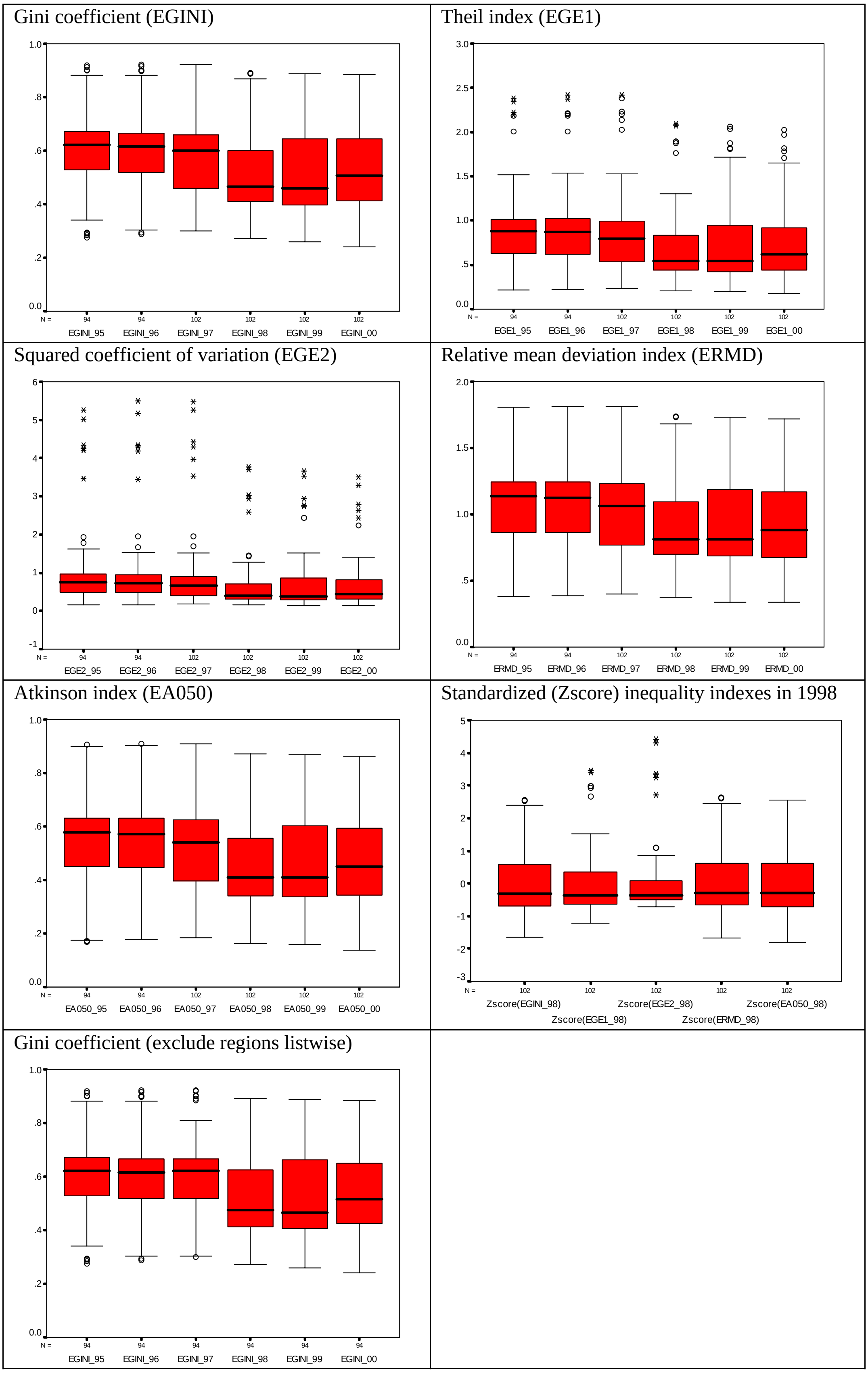


Figure 7, which presents the boxplots of the Gini coefficient on education level completed, shows that the Portuguese regions of Norte, Centro, Alentejo and Algarve are outliers from the upper edge of the box, while educational inequality within Sachsen and Thüringen are between 1.5 and 3 box lengths from the lower edge. The boxplots of the generalised entropy indices (the Theil index and the squared coefficient of variation) show many outliers and extreme regions. For instance, Açores, Madeira, Centro (PT), Alentejo and Algavre depict extreme regions for the Theil index in 1995. Portuguese regions are either outliers or extreme cases. Considering the squared coefficient of variation, there are many extreme regions. Their value is very high and they represent Portuguese regions only. The Spanish region of Centro is also an outlier over the period 1995-1998. The distributions of the relative mean deviation index are less skewed, because two regions are outliers (Açores and Alentejo) in 1998 only. The distributions of the Atkinson index are compact as well. Madeira and Açores are outlying observations at the higher end of the distribution in 1995 and 1996 respectively; and Hamburg, Brandenburg, Sachsen and Sachsen-Anhalt are outlying regions in the lower end of the distribution. Final, for all educational inequality indices, the median and the average have decreased considerably from 1997 to 1998. For instance, the mean and the average of the Gini coefficient decreased by 0.07 and 0.13, respectively.

Figure 7: Boxplot for inequality indices on education level completed



Note: extreme cases and outliers are sorted by descending order:
EGINI: PT3, PT2, PT15 and PT14 (upper end); DE4, DEE, DED and DEG (lower end) in 1995; PT2, PT3, PT12, PT14, PT15 and PT11 (upper end); DED and DEG (lower end) in 1996; PT14 and PT12 (upper end) in 1998.

EGE1: PT3, PT2, PT15, PT14 and PT11 (upper end) in 1995; PT2, PT3, PT12, PT14, PT15 and PT11 (upper end) in 1996; PT2, PT3, PT14, PT12, PT15 and PT11 (upper end) in 1997; PT14, PT2, PT12, PT15, PT3 and PT11 (upper end) in 1998; PT2, PT14, PT12, PT3 and PT15 (upper end) in 1999; PT2, PT14, PT15, PT12 and PT3 (upper end) in 2000.
EGE2: PT3, PT2, PT15, PT14, PT12, PT11, PT13 and ES4 (upper end) in 1995; PT2, PT3, PT12, PT14, PT15, PT11, PT13 and ES4 (upper end) in 1996; PT2, PT3, PT14, PT12, PT15, PT11, PT13 and ES4 (upper end) in 1997; PT14, PT2, PT12, PT15, PT3, PT11, ES4 and PT13 (upper end) in 1998; PT2, PT14, PT12, PT3, PT15 and PT11(upper end) in 1999; PT2, PT14, PT15, PT12, PT3, PT11 (upper end) in 2000.
ERMD: PT14 and PT2 (upper end) in 1998.
EA050: PT3 (upper end); DE4, DED and DEG (lower end) in 1995; PT2 (upper end) in 1996.
EGINI (exclude regions listwise): PT3, PT2, PT15 and PT14 (upper end); DE4, DEE, DED and DEG (lower end) in 1995; PT2, PT3, PT12, PT14, PT15 and PT11 (upper end); DED and DEG (lower end) in 1996; PT2, PT3, PT14, PT12, PT15 and PT11 (upper end); DED (lower end) in 1997 (see Appendix A.1).

The distributions of educational inequality indices are comparable only when they are measured on the same scale. Illustrating the boxplots of the standardised educational inequality indices in 1998 (Figure 7), the distributions of the Gini coefficient, the relative mean deviation index and the Atkinson index are quite similar to each other and are the most compact. The normality assumption is rejected for all distributions, because the ratio of skewness to their standard error is greater than +2 which indicates a long right tail[17]. Table 3 demonstrates the Pearson correlation of the above indices in 1998. Correlations are high and up to 0.861. They are also significant at the 0.01 level (2-tailed) and at the first three decimals.

Table 3: The Pearson correlations among education level completed inequality indices in 1998

	EGINI	EGE1	EGE2	ERMD	EA050
EGINI	1	0.966 (0.000)** 102	0.867 (0.000)** 102	0.985 (0.000)** 102	0.990 (0.000)** 102
EGE1		1	0.963 (0.000)** 102	0.971 (0.000)** 102	0.965 (0.000)** 102
EGE2			1	0.874 (0.000)** 102	0.861 (0.000)** 102
ERMD				1	0.996 (0.000)** 102
EA050					1

Note: ** correlation is significant at the 0.01 level (2-tailed).

The next step is to identify global and local spatial autocorrelation so as to characterise the way inequalities in educational attainment are located in the EU and the way this pattern has probably changed over 1995-2000. Due to the high correlation among inequality indices in education level completed, we only present the spatial autocorrelation analysis for the Gini coefficient. First of all, Moran’s I statistics computed using the rook first order contiguity spatial weights matrices over 1995-2000 show a significant positive spatial autocorrelation (Table 4). This is likely to test the interregional interaction through educational externalities. The space-time correlations are also high. For instance, Moran’s I statistic between a within-region inequality in 2000 and inequality of neighbouring regions in 1998 is 0.6809. Taking into account the 3-nearest neighbours spatial weights schemes, Moran’s I statistics are high over 1995-2000. Final, Moran’s I statistics based on the distance band are much lower than the previous schemes, but remain significant. The trends in evolution of the standardised Moran’s I

[17] The ratio of skewness to standard error is 2.97 for the Gini coefficient, 7.48 for the Theil index, 12.41 for the squared coefficient of variation, 3.47 for the relative mean deviation index and 3.64 for the Atkinson index.

statistics are quite similar. It shows a global significant tendency toward spatial clustering of similar regions in terms of educational inequality.

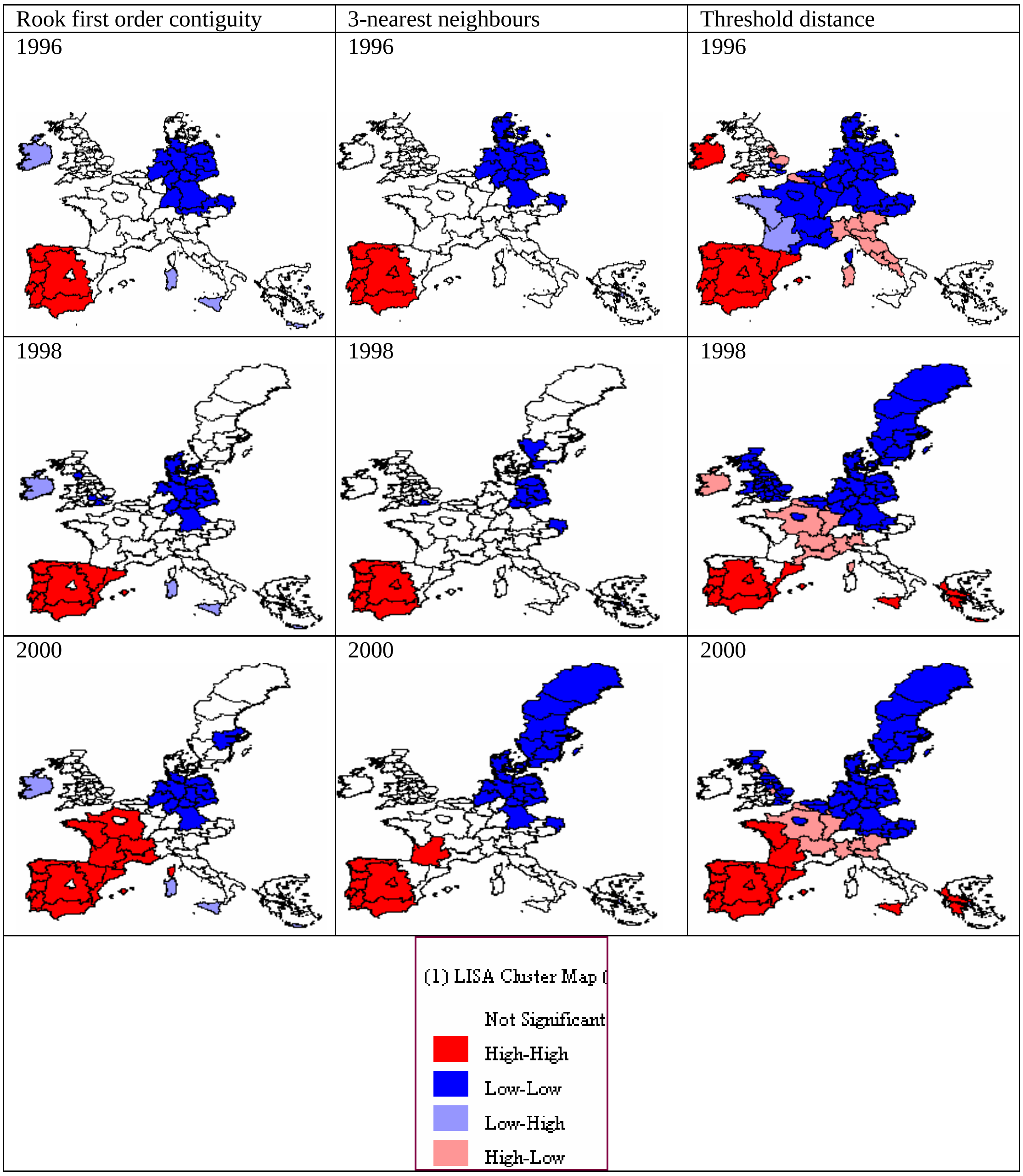
Table 4: Moran's I for the Gini coefficient on education level completed (EGINI)

		13 countries (E[I]=-0.0099)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995												
	1996												
	1997	0.6906	-0.0050	0.0741	9.3873	0.7983	-0.0089	0.0744	10.8495	0.4686	-0.0097	0.0228	20.9781
	1998	0.7063	-0.0090	0.0748	9.5628	0.8217	-0.0076	0.0721	11.5021	0.4643	-0.0101	0.0219	21.6621
	1999	0.7224	-0.0104	0.0741	9.8893	0.7619	-0.0078	0.0742	10.3733	0.3943	-0.0107	0.0216	18.7500
	2000	0.7195	-0.0100	0.0777	9.3887	0.7803	-0.0069	0.0743	10.5949	0.4212	-0.0094	0.0223	19.3094
Space-time correlation	1998												
	2000	0.6809	-0.0070	0.0736	9.3465	0.7716	-0.0084	0.0702	11.1111	0.4301	-0.0102	0.0213	20.6714
		Excluded SE (E[I]=-0.0108)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995	0.7229	-0.0102	0.0769	9.5332	0.8223	-0.0125	0.0767	10.8840	0.3889	-0.0100	0.0223	17.8879
	1996	0.6995	-0.0101	0.0749	9.4740	0.7913	-0.0121	0.0789	10.1825	0.3783	-0.0111	0.0235	16.5702
	1997	0.6764	-0.0107	0.0740	9.2851	0.7730	-0.0102	0.0745	10.5128	0.3892	-0.0115	0.0227	17.6520
	1998	0.7124	-0.0098	0.0756	9.5529	0.8195	-0.0123	0.0782	10.6368	0.4370	-0.0110	0.0229	19.5633
	1999	0.7257	-0.0088	0.0726	10.1171	0.7535	-0.0092	0.0766	9.9569	0.3558	-0.0119	0.0225	16.3422
	2000	0.7204	-0.0069	0.0719	10.1154	0.7692	-0.0135	0.0771	10.1518	0.3632	-0.0107	0.0217	17.2304
Space-time correlation	1998	0.5713	-0.0070	0.0661	8.7489	0.6689	-0.0096	0.0717	9.4630	0.3566	-0.0084	0.0210	17.3810
	2000	0.6843	-0.0068	0.0758	9.1174	0.7653	-0.0075	0.0750	10.3040	0.3922	-0.0102	0.0220	18.2909

Note: All statistics are significant at p=0.001.

Once again, LISA is required so as to compare the human capital inequality values of each specific region with the values in the neighbouring regions. Figure 8 depicts the cluster map for the Gini coefficient on educational inequality in 1996, 1998 and 2000 at three weights schemes. The cluster maps of the first order contiguity schemes and the 3-nearest neighbours ones are quite similar. Portugal and Spain include clusters of regions with high educational inequality, while Germany and Denmark include clusters with low human capital inequality. In 2000, both clusters are expanded even more including some Western French regions (i.e. Sud-Ouest) for the high inequality human capital cluster and some Swedish regions (i.e. Östra Mellansverige) for the low one. Considering the distance band weights schemes, the clusters are evenly spread out and also are separated by a buffer zone which includes at least the regions of Bassin Parisien, Nord – Pas-de-Calais, Est, Centre-Est, Nord Ovest and Lombardia, in 1998 and 2000.

Figure 8: Cluster map for the Gini coefficient on education level completed (EGINI) in 1996, 1998 and 2000



The cluster maps highlight some spatial heterogeneity hidden in the global spatial autocorrelation pattern. This possibly indicates the coexistence of two distinct spatial regimes. Firstly, urbanisation seems to be negatively correlated with human capital inequality, because it is lower in metropolis. Secondly, there is empirical evidence of an EU North-South divide. Homogeneity is higher among the Northern regions of the EU, as well as among the Southern ones, rather than between North-South regions. Although all regions benefit from the human capital diffusion which seems to be easier within groups of closely related economies (Vaya, Lopez-Bazo et al. 2004). Responses to educational inequality changes over the period 1995-2000 remained almost constant demonstrating the persistence of inequality and its dynamic process.

The spatial distribution, the boxplots, the Pearson correlations among inequality indices, the Moran's I statistic and the cluster maps for inequality in age when the highest education level was completed are displayed in Appendices A8-A12.

The spatial distribution of educational inequality within regions when it is measured as inequality in age seems to be different from inequality in education level completed (Appendix A.8). In both cases, however, the geographical distribution appears to be far from random or equal. The Gini coefficient is almost double of the EU average in Northern Italy (Nord Ovest, Lombardia, Nord Est and Emilia-Romagna), in Southern Portugal (Lisboa, Alentejo and Algarve) and in German regions of Brandenburg and Sachsen. Another important characteristic of this figure is the within-country disparities of the Gini coefficient. In Portugal, Spain, Italy and Germany, regional disparities fluctuate at high Gini coefficient levels, while in the United Kingdom and France at low ones. The above argument underlines the importance of the within-country disparities in inequalities considering the broader concept of human capital which is likely to encompass experience, learning-by-doing and on-the-job training from the positive point of view, and unemployment and economic inactivity period from the negative point of view. The geographic distributions of other measures of inequalities such as the relative mean deviation index, the Theil index, the squared coefficient of variation and the Atkinson index yield similar results[18]. Educational inequality hence seems to be concentrated in particular regions of the EU.

Considering boxplots (Appendix A.9), all distributions are fairly compact, because the whiskers are in fact the extreme values. The interquartile range seems to be constant between 1995 and 1997, and between 1998 and 2000. The boxplots of standardised education inequality indices in 1998 demonstrate that the distributions of the Gini coefficient and the relative mean deviation index have the greatest difference between the first and third quartiles. Additionally, their compactness is similar to each other. The normality assumption is accepted for the Gini coefficient, the relative mean deviation and the Atkinson distribution, because the ratio of skewness to their standard error is greater than -2 and less than +2, but it is rejected for the generalised entropy indexes (the Theil index and the squared coefficient of variation)[19]. Correlations among these indices in 1998 are high and up to 0.94 (Appendix A.10). They are also significant at the 0.01 level (2-tailed) and at the first three decimals. Moran's I

---

[18] The results are provided under request.

[19] The ratio of skewness to standard error is 0.31 for the Gini coefficient, 2.02 for the Theil index, 2.36 for the squared coefficient of variation, 0.32 for the relative mean deviation index and 1.88 for the Atkinson index. Following the sensitivity of the Atkinson index for income, once again, this index should become more sensitive to 'transfers' among people who completed their highest formal studies when they were young and less sensitive to 'transfers' among people who completed their studies when they were older. Additionally at higher values sensitivity parameters of the Atkinson index (i.e. $\varepsilon = 1$ and $\varepsilon = 2$), it fits better to the normal distribution because the ration of skewness to standard error is lower (i.e., AA100=1.72 and AA200=1.42, respectively).

statistic for any spatial weights schemes over the period 1995-2000 depicts a positive spatial autocorrelation (Appendix A.11). The space-time statistics are also high. The cluster maps in 1996, 1998 and 2000 confirm the local variation in the spatial autocorrelation (Appendix A.12). Inequality in human capital is concentrated in particular areas of Europe. The regions with relatively high educational inequality (respectively low) are localised close to other regions with relatively high educational inequality (respectively high) more often than their localisations were purely random. Different trends in inequality distribution exist over the EU space. The weights schemes of the first order contiguity and the 3-nearest neighbours show that clusters of regions with high educational inequality are in Central and Northern Italy (Nord Ovest, Lombardia, Emilia-Romagna and Centro), in Southern Portugal (Lisboa, Alentejo and Algavre) and in Eastern Germany (Brandenburg, Mecklenburg-Vorpommern and Sachsen). Additionally, Southern Portugal cluster is even more expanded for the distance band schemes including Southern Spanish regions. In contrast, most British regions are clusters with low human capital inequality. The distance band weights schemes create bigger clusters than the previous schemes. Furthermore, Noroeste, Noreste, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz and Saarland are spatial outliers over 1996-2000. The maps and the boxplots depict one source of spatial heterogeneity which is the urbanisation. It seems to be negatively correlated with educational inequality. The figures show an increase in the homogeneity within urban centers and within rural areas. Spatial autocorrelation seems to favour the diffusion of human capital activities from an urban center to another or from the inner to the outer city, rather than from the urban to the periphery. Nevertheless, the distance between these groups remained the same underlining the stagnation of the polarisation or stratification process, on the one hand, and the persistence of educational inequality, on the other. Hence the existence of autocorrelation, heteroskedasticity and persistence highlight the requirement for space-time analysis on educational inequality.

The relationship between inequality in education level completed and inequality in age when the highest education level was completed is investigated by the comparison of their boxplots, the Pearson correlation index and the bivariate Moran's I statistic. First, in view of the boxplots of proxies for educational inequality (Appendix A.13), it is shown that the distributions of both proxies are quite compact. However, the difference between the two whiskers of inequality in education level completed distributions is approximately triple of that inequality in age. Furthermore, the minimum value of the inequality in education level completed distribution is the maximum one of the other proxy for educational inequality. Second, measuring the Pearson correlation index, a positive linear correlation is illustrated in 1998, 1999 and 2000 (Appendix A.14). Third, the correlation between the Gini coefficient on education level completed at a region and the Gini coefficient on age when the highest education level was completed at neighbouring regions, and vice versa, are explored. In 1996, for instance, the bivariate Moran's I statistic is not significant as well. Conversely, in 1998, where the Pearson correlation has the highest value, the bivariate Moran's I statistic between inequality in education level completed at a region and inequality in age of neighbouring regions is 0.3712, while between inequality in age at a region and inequality in education level completed of neighbouring regions is 0.3843 for first order contiguity spatial weights schemes, 0.4663 and 0.4323, respectively, for 3-nearest neighbours weights schemes, and 0.2742 and 0.2797, respectively, for threshold distance band weights schemes. These statistics are significant at the 0.001 level.

5. Correlation between educational attainment and inequality

Table 5 illustrates the Pearson correlations between the average and inequality in education level completed. The relationship is negative and statistically significant at the 0.01 level. The higher the educational attainment, the lower the educational inequality, and vice versa. Education seems to be one of the most powerful instruments known for reducing educational inequality. The increased opportunity of acquiring higher education is likely to reduce the educational inequality as more people are enable to improve their socioeconomic circumstances. Educational expansion seems to offer educational opportunities and numerous favourable chances to both advantaged and disadvantaged groups.

Table 5: Pearson correlation between average education level completed (EMN) and inequality in education level completed

	1995	1996	1997	1998	1999	2000
EGINI	-0.899 (0.000)** 94	-0.869 (0.000)** 94	-0.892 (0.000)** 94 -0.901 (0.000)** 102	-0.900 (0.000)** 94 -0.9 (0.000)** 102	-0.902 (0.000)** 94 -0.902 (0.000)** 102	-0.880 (0.000)** 94 -0.882 (0.000)** 102
EGE1	-0.898 (0.000)** 94	-0.876 (0.000)** 94	-0.890 (0.000)** 94 -0.898 (0.000)** 102	-0.853 (0.000)** 94 -0.854 (0.000)** 102	-0.866 (0.000)** 94 -0.866 (0.000)** 102	-0.877 (0.000)** 94 -0.879 (0.000)** 102
EGE2	-0.785 (0.000)** 94	-0.771 (0.000)** 94	-0.781 (0.000)** 94 -0.784 (0.000)** 102	-0.766 (0.000)** 94 -0.769 (0.000)** 102	-0.791 (0.000)** 94 -0.794 (0.000)** 102	-0.811 (0.000)** 94 -0.815 (0.000)** 102
ERMD	-0.904 (0.000)** 94	-0.876 (0.000)** 94	-0.894 (0.000)** 94 -0.902 (0.000)** 102	-0.856 (0.000)** 94 -0.849 (0.000)** 102	-0.858 (0.000)** 94 -0.852 (0.000)** 102	-0.855 (0.000)** 94 -0.853 (0.000)** 102
EA050	-0.907 (0.000)** 94	-0.879 (0.000)** 94	-0.896 (0.000)** 94 -0.905 (0.000)** 102	-0.869 (0.000)** 94 -0.865 (0.000)** 102	-0.871 (0.000)** 94 -0.868 (0.000)** 102	-0.860 (0.000)** 94 -0.862 (0.000)** 102

Note: ** correlation is significant at the 0.01 level (2-tailed).

Table 6 shows the Pearson correlations between the average and inequality in age when the highest education level was completed. This relationship is positive but not statistically significant for the squared coefficient of variation over the period 1995-2000 and for the Theil and the Atkinson index between 1995 and 1997. This is probable because occupations that require high levels of investment in human capital are beyond the reach of poor people, who choose instead to work for other (Banerjee and Newman 1991, 1993). Another possibly explanation is that the poor require relatively higher returns to increase expenditure on education, so they in vest in education smaller shares of their income than the rich (Ceroni 2001). Encompassing experience, learning by doing and on-the-job training may positively affect educational inequality as they are likely to offer opportunities to already advantaged groups only. For instance, people with more work experience may be more informed and make better choices than those with little experience.

Table 6: Pearson correlation between average in age when the highest education level was completed (AMN) and inequality in age when the highest education level was completed

	1995	1996	1997	1998	1999	2000
AGINI	0.240 (0.027)* 85 0.267 (0.010)** 93	0.238 (0.028)* 85 0.263 (0.011)* 93	0.251 (0.020)* 85 0.275 (0.008)** 93	0.360 (0.001)** 85	0.356 (0.001)** 85	0.351 (0.001)** 85
AGE1	0.125 (0.254) 85 0.165 (0.114) 93	0.130 (0.236) 85 0.166 (0.112) 93	0.140 (0.200) 85 0.174 (0.095) 93	0.265 (0.014)* 85	0.257 (0.017)* 85	0.243 (0.025)* 85
AGE2	0.064 (0.559) 85 0.110 (0.294) 93	0.071 (0.517) 85 0.113 (0.283) 93	0.077 (0.481) 85 0.117 (0.265) 93	0.207 (0.058) 85	0.210 (0.053) 85	0.177 (0.105) 85
ARMD	0.261 (0.016)* 85 0.288 (0.005)** 93	0.265 (0.014)* 85 0.290 (0.005)** 93	0.282 (0.009)** 85 0.305 (0.003)** 93	0.389 (0.000)** 85	0.375 (0.000)** 85	0.364 (0.001)** 85
AA050	0.147 (0.179) 85 0.184 (0.077) 93	0.151 (0.168) 85 0.185 (0.077) 93	0.163 (0.136) 85 0.195 (0.062) 93	0.285 (0.008)** 85	0.273 (0.012)* 85	0.268 (0.013)* 85

Note: ** correlation is significant at the 0.01 level (2-tailed); * correlation is significant at the 0.05 level (2-tailed).

6. Conclusion

The European regions differ with regard to human capital endowment. The geographical distribution of educational attainment and inequality is not uniform. It is characterised by significant positive global spatial autocorrelation and space-time correlation. The evolution of education within a region is closely related to the evolution of neighbouring regions (denoting spatial autocorrelation). The spatial evolution of education affects the dynamic evolution of human capital through geographical distances and proximity (showing space-time correlation). Positive spatial dependence shows a region surrounded by high-educated economies can achieve higher educational stock. The reverse is also true. Moran's I statistics lead to the same results for the sign (positive) and significance of global spatial dependence, highlighting the robustness of the results, with regard to the choice of the spatial weights matrix. Since labour is a mobile production factor, public infrastructure investments in one region can draw production away from other regions or provide access to adjacent regions not previously accessible (Lall and Yilmaz 2001). Regional variations in educational attainment and inequality are likely to depict regional variations in the average and inequality in skills, efforts, opportunities, knowledge and aspiration, on the one hand; and national institutional differences, on the other. The application of the global and local spatial association tests permits the detection of educational patterns in the territory of the EU which has not been changed dramatically throughout the whole period of study, denoting the persistence of educational attainment and inequality in specific regions. Human capital is an important factor on shaping regional interactions. Regional disparities in education are influenced by region and national specific characteristics (i.e. educational guidelines) and the availability of high-educated people in neighbouring regions.

The ESDA on education stresses some kind of spatial heterogeneity hidden in the spatial autocorrelation pattern. Spatial effects perform differently in two regimes: the urbanisation pattern and the European North-South one. There are systematic differences between urban and rural European regions and between North and South European regions. Educational attainment is higher in the North and in urban areas, while educational inequality is lower in these areas. Because of spatial interactions between regions, geographical location (urban or rural, and North or South) is important in accounting for human capital performance of regions. Regions are geographically correlated due to some processes, which connect different areas, like educational diffusion, and national institutions. Vaya, Lopez-Bazo et al. (2004: 433) point out that externalities spill over the barriers of regional economies, in line with the idea of across economy interactions outlined in Lucas (1988; 1993). They also mention that there are spatial limits to the spread of externalities and the diffusion of skills and knowledge will always be easier within groups of closely related economies (‘clubs’). Economies within a group (i.e. group of Northern European countries) interact more with each other than with those outside. The diffusion of human capital seems to be even stronger between regions of the same economy than the diffusion between national economies. The analysis shows that educational policies should account for the spillover effects with adjoining regions. The prevalence of interregional educational externalities may have created a ‘poverty human capital trap’ based on geographical location.

Finally, the within-region component of educational inequality constitutes the major portion of the European inequality, while the between-region component represents the minor portion.

Bibliography

Akita, T. (2003). "Decomposing Regional Income Inequality in China and Indonesia Using Two-Stage Nested Theil Decomposition Method." Annals of Regional Science **37**(1): 55-77.

Amrhein, C. G. (1995). "Searching for the Elusive Aggregation Effect - Evidence from Statistical Simulations." Environment and Planning A **27**(1): 105-119.

Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht; Boston, Kluwer Academic Publishers.

Anselin, L. (1995). "Local Indicators of Spatial Association - LISA." Geographical Analysis **27**(2): 93-115.

Anselin, L. (2003a). GeoDa™ 0.9 User’s Guide, Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign.

Anselin, L. (2003b). An Introduction to Spatial Autocorrelation Analysis with GeoDa, Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign.

Anselin, L. and A. K. Bera (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. Handbook of Applied Economic Statistics. A. Ullah and D. E. A. Giles. New York, Marcel Dekker**:** 237-289.

Anselin, L., R. J. G. M. Florax, et al., Eds. (2004). Advances in Spatial Econometrics: Methodology, Tools and Applications. Advances in spatial science. Berlin, Springer.

Arbia, G. (1989). Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems. Dordrecht, Kluwer Academic.

Arthur, W. B. (1994). Increasing Returns and Path Dependence in the Economy. Ann Arbor,, c1994.

Banerjee, A. V. and A. F. Newman (1991). "Risk-Bearing and the Theory of Income Distribution." Review of Economic Studies **58**(2): 211-235.

Banerjee, A. V. and A. F. Newman (1993). "Occupational Choice and the Process of Development." Journal of Political Economy **101**(2): 274-298.

Barro, R. J. (1991). "Economic Growth in a Cross-Section of Countries." Quarterly Journal of Economics **106**(2): 407-443.

Barro, R. J. (2001). "Human Capital and Growth." American Economic Review **91**(2): 12-17.

Baumont, C., C. Ertur, et al. (2003). Spatial Convergence Clubs and the European Regional Growth Process. European Regional Growth. B. Fingleton. Berlin, Springer**:** 131–158.

Becker, G. S. and B. R. Chiswick (1966). "Economics of Education and Distribution of Earnings." American Economic Review **56**(2): 358-369.

Benhabib, J. and M. M. Spiegel (1994). "The Role of Human Capital in Economic Development Evidence from Aggregated Cross-Country Data." Journal of Monetary Economics **34**(2): 143-173.

Bennett, R., H. Glennerster, et al. (1995). "Regional Rates of Return to Education and Training in Britain." Regional Studies **29**(3): 279-295.

Castello, A. and R. Domenech (2002). "Human Capital Inequality and Economic Growth: Some New Evidence." Economic Journal **112**(478): C187-C200.

Centre for Educational Research and Innovation and Organisation for Economic Co-operation and Development (1998). Human Capital Investment: an International Comparison. Paris, Organisation for Economic Co-operation and Development.

Ceroni, C. B. (2001). "Poverty Traps and Human Capital Accumulation." Economica **68**(270): 203-219.

Ciccone, A. (2004). Human Capital as a Factor of Growth and Employment at the Regional Level: the case of Italy, Report for the European Commission, DG for Employment and Social Affairs.

Cornia, G. A., T. Addison, et al. (2001). Income Distribution Changes and their Impact in the Post-World War II Period, WIDER Discussion Paper No. 2001/89, World Institute for Development Economics Research, United Nations University.

Cressie, N. A. C. (1993). Statistics for Spatial Data. New York, Wiley.

Davanzo, J. (1976). "Differences between Return and Nonreturn Migration - Econometric Analysis." International Migration Review **10**(1): 13-27.

Davanzo, J. (1978). "Does Unemployment Affect Migration - Evidence from Micro Data." Review of Economics and Statistics **60**(4): 504-514.

Davis, D. R. and D. E. Weinstein (2003). "Market Access, Economic Geography and Comparative Advantage: an Empirical Test." Journal of International Economics **59**(1): 1-23.

De Gregorio, J. and J. W. Lee (2002). "Education and Income Inequality: New Evidence from Cross-Country Data." Review of Income and Wealth(3): 395-416.

Durlauf, S. N. and P. A. Johnson (1995). "Multiple Regimes and Cross-Country Growth Behavior." Journal of Applied Econometrics **10**(4): 365-384.

Ertur, C. and J. Le Gallo (2003). An Exploratory Spatial Data Analysis of European Regional Disparities, 1980-1995. European Regional Growth. B. Fingleton. Berlin, Springer**:** 55-98.

Fingleton, B., Ed. (2003). European Regional Growth. Advances in spatial science. Berlin, Springer.

Fischer, M. M. and C. Stirbock (2006). "Pan-European Regional Income Growth and Club-Convergence." Annals of Regional Science **40**(4): 693-721.

Florax, R. J. G. M. and S. Rey (1995). The Impacts of Misspecified Spatial Interaction in Linear Regression Models. New Directions in Spatial Econometrics. L. Anselin and R. J. G. M. Florax. Berlin; New York, Springer**:** 111-135.

Fotheringham, A. S., C. Brunsdon, et al. (2000). Quantitative Geography: Perspectives on Spatial Data Analysis. London, SAGE.

Fotheringham, A. S., C. Brunsdon, et al. (2002). Geographically Weighted Regression: the Analysis of Spatially Varying Relationships, Wiley.

Fujita, M., P. R. Krugman, et al. (1999). The Spatial Economy: Cities, Regions and International Trade. Cambridge, Mass., MIT Press.

Fujita, M. and J.-F. Thisse (2002). Economics of Agglomeration: Cities, Industrial Location, and Regional Growth. Cambridge, Cambridge University Press.

Gemmell, N. (1996). "Evaluating the Impacts of Human Capital Stocks and Accumulation on Economic Growth: Some New Evidence." Oxford Bulletin of Economics and Statistics **58**(1): 9-&.

Gujarati, D. N. (2003). Basic Econometrics. Boston, London, McGraw-Hill.

Haining, R. P. (1995). Data Problems in Spatial Econometric Modeling. New Directions in Spatial Econometrics. L. Anselin and R. J. G. M. Florax. Berlin; New York, Springer**:** 156-171.

Hansen, M. N. (2001). "Education and Economic Rewards. Variations by Social-Class Origin and Income Measures." European Sociological Review **17**(3): 209-231.

Hanushek, E. A. and D. D. Kimko (2000). "Schooling, Labor-Force Quality, and the Growth of Nations." American Economic Review **90**(5): 1184-1208.

Helpman, E. and P. R. Krugman (1985). Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy. Brighton, Wheatsheaf.

Hirschman, A. O. (1958). The Strategy of Economic Development. New Haven, Yale University Press.

Kaldor, N. (1970). "Case for Regional Policies." Scottish Journal of Political Economy **17**(3): 337-348.

Kaldor, N. (1981). "The Role of Increasing Returns, Technical Progress and Cumulative Causation in the Theory of International Trade and Economic Growth." Economie Applique **XXXIV**: 593-617.

Kaldor, N. (1985). Economics without Equilibrium. Cardiff, University College Cardiff Press.

Krugman, P. (1980). "Scale Economies, Product Differentiation, and the Pattern of Trade." American Economic Review **70**(5): 950-959.

Krugman, P. (1993). "On the Number and Location of Cities." European Economic Review **37**(2-3): 293-298.

Krugman, P. and A. J. Venables (1995). "Globalization and the Inequality of Nations." Quarterly Journal of Economics **110**(4): 857-880.

Krugman, P. and A. J. Venables (1996). "Integration, Specialization, and Adjustment." European Economic Review **40**(3-5): 959-967.

Krugman, P. R. (1991a). Geography and Trade. Cambridge; Leuven, Belgium, MIT Press: Leuven University Press.

Krugman, P. R. (1991b). "Increasing Returns and Economic Geography." Journal of Political Economy **99**(3): 489-499.

Lall, S. V. and S. Yilmaz (2001). "Regional Economic Convergence: Do Policy Instruments Make a Difference?" Annals of Regional Science **35**(1): 153-166.

Lucas, R. E. (1988). "On the Mechanics of Economic Development." Journal of Monetary Economics **22**(1): 3-42.

Lucas, R. E. (1993). "Making a Miracle." Econometrica **61**(2): 251-272.

Manski, C. F. (1993). "Identification of Endogenous Social Effects - the Reflection Problem." Review of Economic Studies **60**(3): 531-542.

Marin, A. and G. Psacharopoulos (1976). "Schooling and Income Distribution." Review of Economics and Statistics **58**(3): 332-338.

Martin, P. (1998). "Can Regional Policies Affect Growth and Geography in Europe?" World Economy **21**(6): 757-774.

Martin, P. (1999a). "Are European Regional Policies Delivering?" European Investment Bank Papers **4**(2): 10-23.

Martin, P. (1999b). "Public Policies, Regional Inequalities and Growth." Journal of Public Economics **73**(1): 85-105.

Martin, R. (1999c). "The New 'Geographical Turn' in Economics: Some Critical Reflections." Cambridge Journal of Economics **23**(1): 65-91.

McCann, P. and S. Sheppard (2001). Public Investment and Regional Labour Markets: the Role of UK Higher Education. Public Investment and Regional Economic Development. D. Felzenstein. Cheltenham, UK, Northampton, MA, Edward Elgar**:** 135-153.

Moran, P. A. P. (1950). "Notes on Continuous Stochastic Phenomena." Biometrika **37**(1/2): 17-23.

Murphy, K. M., A. Shleifer, et al. (1991). "The Allocation of Talent - Implications for Growth." Quarterly Journal of Economics **106**(2): 503-530.

Myrdal, G. (1957). Economic Theory and Under-Developed Regions. London, Duckworth.

Openshaw, S. (1983). The Modifiable Areal Unit Problem. Norwick [Norfolk], Geo Books.

Perroux, F. (1950). "Economic Space: Theory and Applications." The Quarterly Journal of Economics **64**(1): 89-104.

Pritchett, L. (1996). Where Has All the Education Gone? Washington, DC, World Bank, Policy Research Dept., Poverty and Human Resources Division.

Psacharopoulos, G. and A.-M. Arriagada (1986). The Educational Attainment of the Labor Force: an International Comparison, World Bank Discussion Paper, Washington, D.C.

Puga, D. (1999). "The Rise and Fall of Regional Inequalities." European Economic Review **43**(2): 303-334.

Puga, D. and A. J. Venables (1996). "The Spread of Industry: Spatial Agglomeration in Economic Development." Journal of the Japanese and International Economies **10**(4): 440-464.

Ram, R. (1990). "Educational Expansion and Schooling Inequality - International Evidence and Some Implications." Review of Economics and Statistics **72**(2): 266-273.

Rodríguez-Pose, A. and M. Vilalta-Bufi (2005). "Education, Migration, and Job Satisfaction: the Regional Returns of Human Capital in the EU." Journal of Economic Geography **5**(5): 545-566.

Rosenstein-Rodan, P. (1943). "The Problem of Industrialization of Eastern and South-Eastern Europe." Economic Journal **53**: 201-211.

Sianesi, B. and J. Van Reenen (2003). "The Returns to Education: Macroeconomics." Journal of Economic Surveys **17**(2): 157-200.

Stein, M. L. (1999). Interpolation of Spatial Data: Some Theory for Kriging. Berlin, Springer.

Tallman, E. W. and P. Wang (1994). "Human Capital and Endogenous Growth Evidence from Taiwan." Journal of Monetary Economics **34**(1): 101-124.

Temple, J. (1999). "A Positive Effect of Human Capital on Growth." Economics Letters **65**(1): 131-134.

Thomas, V., Y. Wang, et al. (2001). Measuring Education Inequality: GINI Coefficients of Education. Washington, D.C., World Bank.

Unwin, A. and D. Unwin (1998). "Exploratory Spatial Data Analysis with Local Statistics." Journal of the Royal Statistical Society Series D-the Statistician **47**: 415-421.

Vaya, E., E. Lopez-Bazo, et al. (2004). Growth and Externalities across Economies: an Empirical Analysis Using Spatial Econometrics. Advances in Spatial Econometrics: Methodology, Tools and Applications. L. Anselin, R. J. G. M. Florax and S. J. Rey. Berlin, Springer**:** 433-455.

Winegarden, C. R. (1979). "Schooling and Income Distribution - Evidence from International Data." Economica **46**(181): 83-87.

Appendix A.1: Regions: code and name

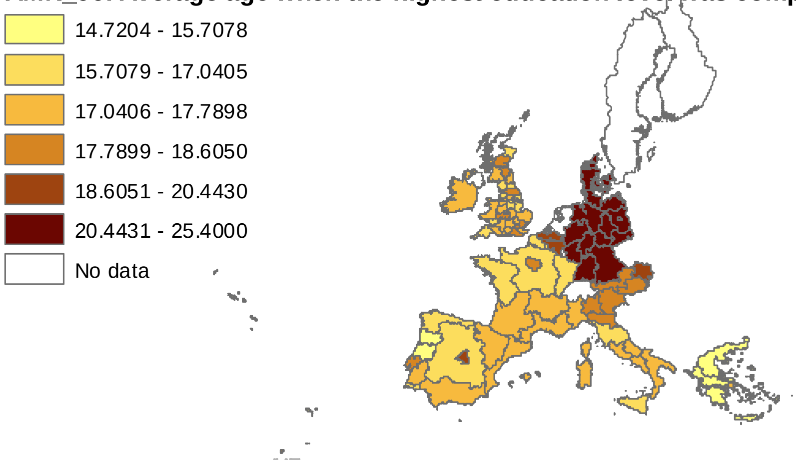
	MICRO-DATA (based on NUTS, version 1995)	
NUTS	CODE	NAME
NUTS0	be	Belgium
NUTS1	be1	Région Bruxelles-capitale/Brussels hoofdstad gewest
NUTS1	be2	Vlaams Gewest
NUTS1	be3	Région Wallonne
NUTS0-NUTS1	dk	Denmark
NUTS0	de	Federal Republic of Germany (including ex-GDR from 1991)
NUTS1	de1	Baden-Württemberg
NUTS1	de2	Bayern
NUTS1	de3	Berlin
NUTS1	de4	Brandenburg
NUTS1	de5	Bremen
NUTS1	de6	Hamburg
NUTS1	de7	Hessen
NUTS1	de8	Mecklenburg-Vorpommern
NUTS1	de9	Niedersachsen
NUTS1	dea	Nordrhein-Westfalen
NUTSNEW	dex	Rheinland-Pfalz+Saarland
NUTS1	ded	Sachsen
NUTS1	dee	Sachsen-Anhalt
NUTS1	def	Schleswig-Holstein
NUTS1	deg	Thüringen
NUTS0	gr	Greece
NUTS1	gr1	Voreia Ellada
NUTS1	gr2	Kentriki Ellada
NUTS1	gr3	Attiki
NUTS1	gr4	Nisia Aigaiou, Kriti
NUTS0	es	Spain
NUTS1	es1	Noroeste
NUTS1	es2	Noreste
NUTS1	es3	Comunidad de Madrid
NUTS1	es4	Centro (ES)
NUTS1	es5	Este
NUTS1	es6	Sur
NUTS1	es7	Canarias (ES)
NUTS0	fr	France
NUTS1	fr1	Île de France
NUTS1	fr2	Bassin Parisien
NUTS1	fr3	Nord - Pas-de-Calais
NUTS1	fr4	Est
NUTS1	fr5	Ouest
NUTS1	fr6	Sud-Ouest
NUTS1	fr7	Centre-Est
NUTS1	fr8	Méditerranée
NUTS0-NUTS1	ie	Ireland
NUTS0	it	Italy
NUTS1	it1	Nord Ovest
NUTS1	it2	Lombardia
NUTS1	it3	Nord Est

NUTS1	it4	Emilia-Romagna
NUTS1	it5	Centro (I)
NUTS1	it6	Lazio
NUTS1	it7	Abruzzo-Molise
NUTS1	it8	Campania
NUTS1	it9	Sud
NUTS1	ita	Sicilia
NUTS1	itb	Sardegna
NUTS0-NUTS1	lu	Luxembourg
NUTS0	at	Austria
NUTS1	at1	Ostösterreich
NUTS1	at2	Südösterreich
NUTS1	at3	Westösterreich
NUTS0	pt	Portugal
NUTS2	pt11	Norte
NUTS2	pt12	Centro (PT)
NUTS2	pt13	Lisboa e Vale do Tejo
NUTS2	pt14	Alentejo
NUTS2	pt15	Algarve
NUTS2	pt2	Açores (PT)
NUTS2	pt3	Madeira (PT)
NUTS0	se	Sweden
NUTS2	se01	Stockholm
NUTS2	se02	Östra Mellansverige
NUTS2	se04	Sydsverige
NUTS2	se06	Norra Mellansverige
NUTS2	se07	Mellersta Norrland
NUTS2	se08	Övre Norrland
NUTS2	se03	Småland med öarna
NUTS2	se05	Västsverige
NUTS0	uk	United Kingdom
NUTS2	uk11	Cleveland, Durham
NUTS2	uk13	Northumberland, Tyne and Wear
NUTS2	uk12	Cumbria
NUTS2	uk81	Cheshire
NUTS2	uk82	Greater Manchester
NUTS2	uk83	Lancashire
NUTS2	uk84	Merseyside
NUTS2	uk21	Humberside
NUTS2	uk22	North Yorkshire
NUTS2	uk23	South Yorkshire
NUTS2	uk24	West Yorkshire
NUTS2	uk31	Derbyshire, Nottinghamshire
NUTS2	uk32	Leicestershire, Northamptonshire
NUTS2	uk33	Lincolnshire
NUTS2	uk71	Hereford and Worcester, Warwickshire
NUTS2	uk72	Shropshire, Staffordshire
NUTS2	uk73	West Midlands (County)
NUTS2	uk40	East Anglia

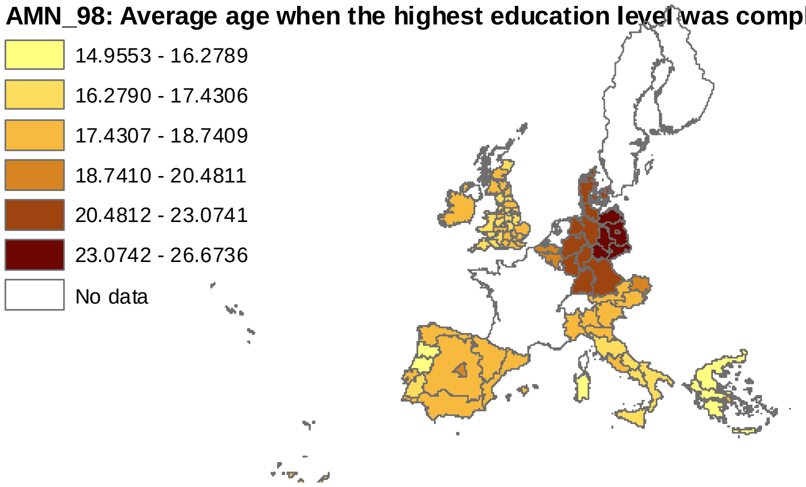
NUTS2	uk51	Bedfordshire, Hertfordshire
NUTS2	uk54	Essex
NUTS2	uk55	Greater London
NUTS2	uk52	Berkshire, Buckinghamshire, Oxfordshire
NUTS2	uk53	Surrey, East-West Sussex
NUTS2	uk56	Hampshire, Isle of Wight
NUTS2	uk57	Kent
NUTS2	uk61	Avon, Gloucestershire, Wiltshire
NUTS2	uk63	Dorset, Somerset
NUTS2	uk62	Cornwall, Devon
NUTS2	uk92	Gwent, Mid-South-West Glamorgan
NUTS2	uk91	Clwyd, Dyfed, Gwynedd, Powys
NUTS2	uka4	Grampian
NUTS2	uka1	Borders-Central-Fife-Lothian-Tayside
NUTS2	uka2	Dumfries and Galloway, Strathclyde

Appendix A.2: Spatial distribution of average age when the highest education level was completed (AMN) in 1996, 1998 and 2000

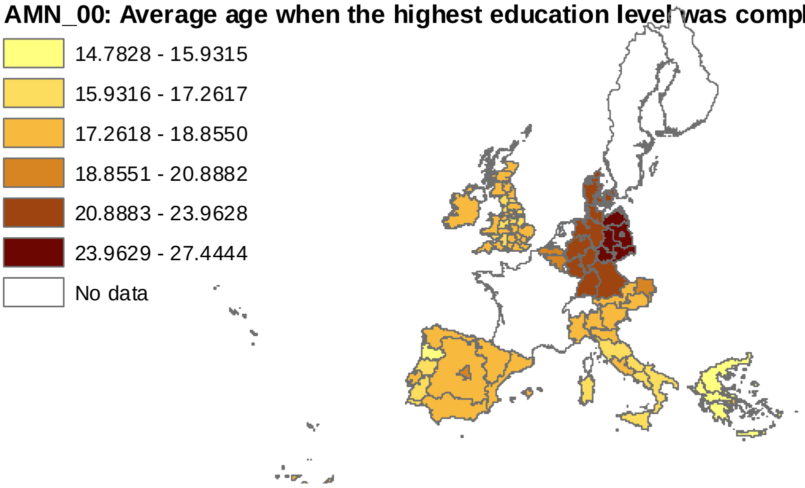
AMN_96: Average age when the highest education level was completed in 1996



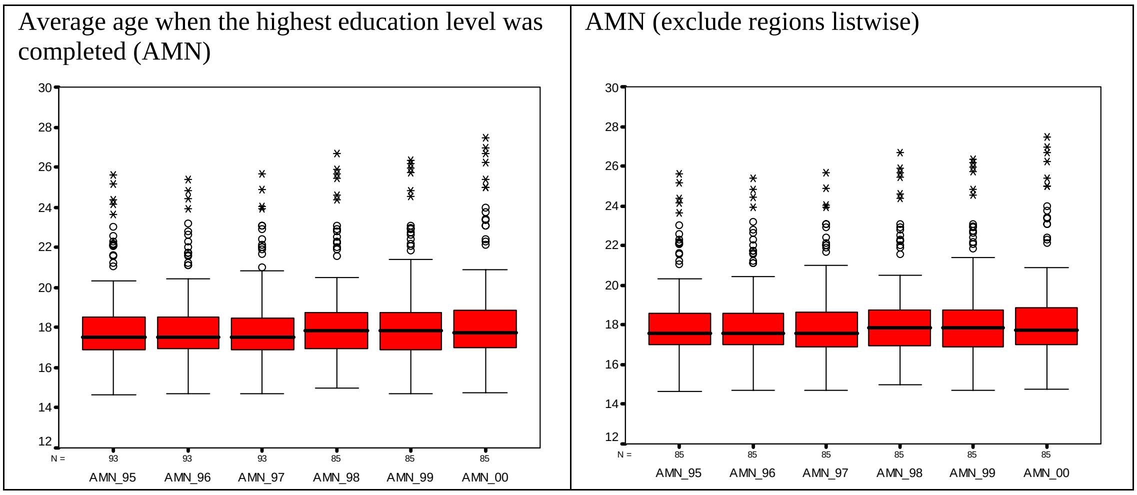
AMN_98: Average age when the highest education level was completed in 1998



AMN_00: Average age when the highest education level was completed in 2000



Appendix A.3: Boxplot for average age when the highest education level completed (AMN)



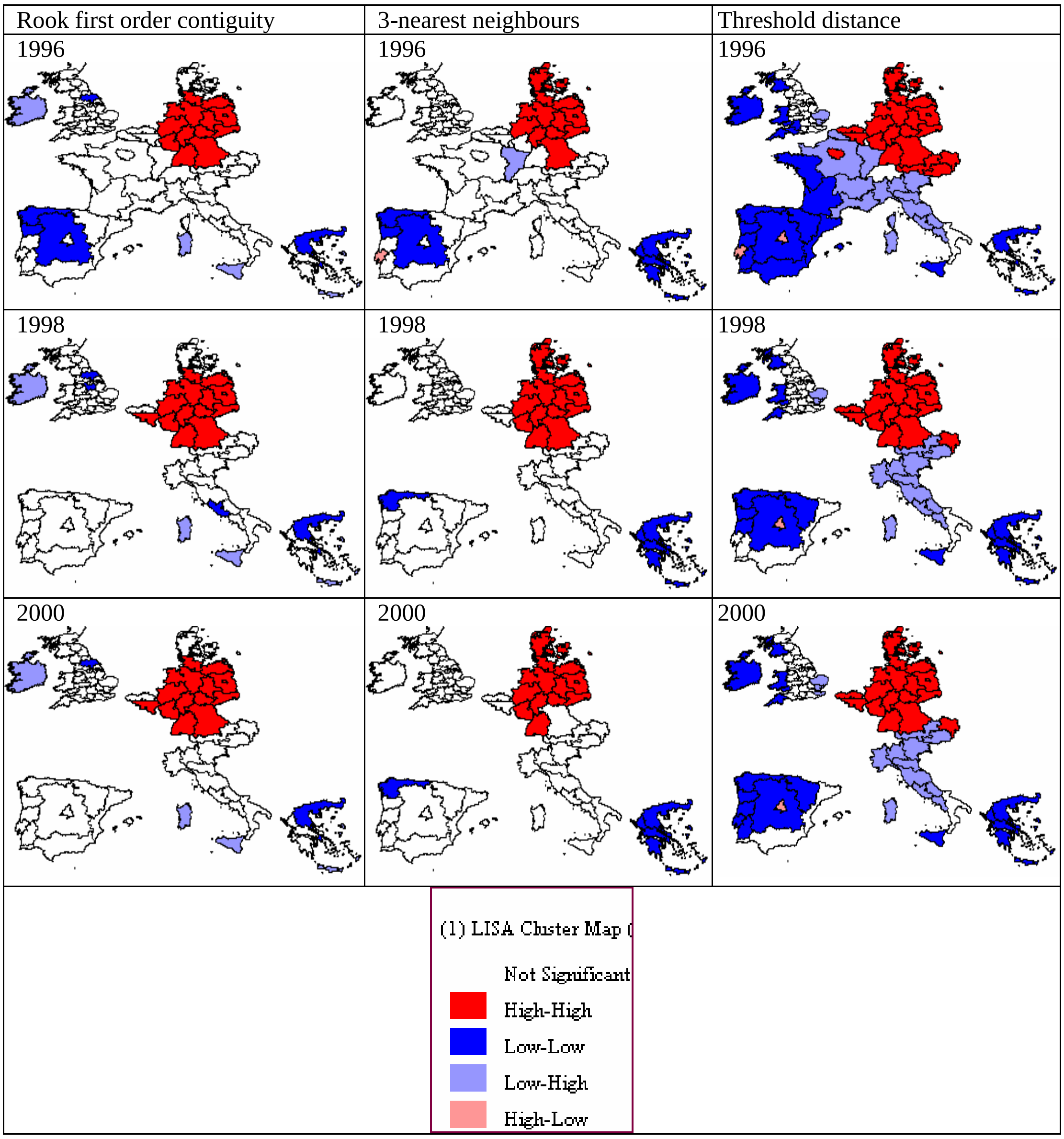
Note: extreme cases and outliers are sorted by descending order: DE3, DED, DE8, DE4, DEE, DEG, DK, DE9, DEA, DE5, DE7, DEF, DEG and DEX in 1995; DE3, DED, DE8, DE4, DEE, DEG, DK, DE9, DEA, DEF, DEF, DE2, DE5, DE6 and DEX in 1996; DE3, DED, DE8, DE4, DEE, DEG, DEF, DEK, DE9, DE5, DEA, DE7, DE2 and DEX (it is not outlier in exclude regions listwise) in 1997; DE3, DED, DE8, DE4, DEG, DEE, DE5, DE9, DEA, DE7, DEF, DE2, DEX, DE6, DK and DE1 in 1998; DED, DE3, DE8, DE4, DEE, DEG, DEF, DEA, DE5, DE7, DE9, DE2, DEX, DEK and DE6 in 1999; DED, DE3, DE8, DE4, DEE, DE5, DEG, DEA, DEF, DE2, DE9, DEX, DE7, DE6, DEK and DE1 in 2000 (see Appendix A.1).

Appendix A.4: Moran's I for average age when the highest education level was completed (AMN)

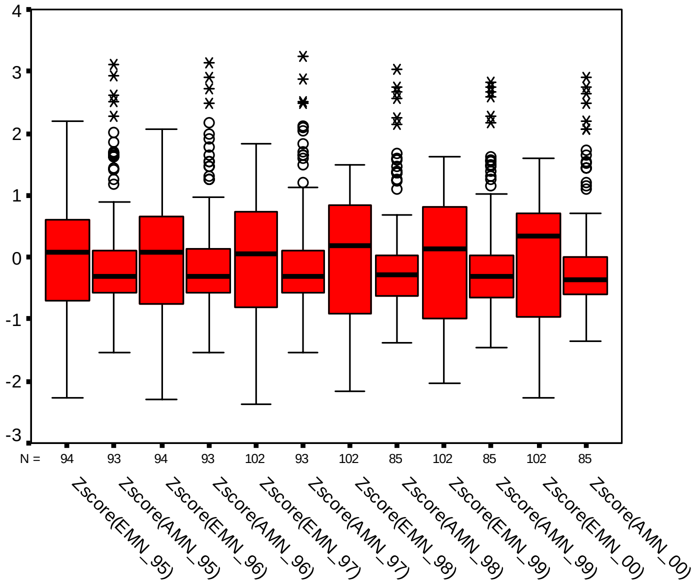
		Excluded SE LU (E[I]=-0.0109)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995	0.7812	-0.0083	0.0764	10.3338	0.8378	-0.0145	0.0725	11.7559	0.2465	-0.0105	0.0223	11.5247
	1996	0.7770	-0.0124	0.0763	10.3460	0.8313	-0.0120	0.0773	10.9094	0.2486	-0.0104	0.0238	10.8824
	1997	0.7872	-0.0140	0.0723	11.0816	0.8365	-0.0126	0.0763	11.1284	0.2464	-0.0105	0.0231	11.1212
	1998												
	1999												
	2000												
Space-time correlation	1998												
	2000												
		Excluded SE LU FR (E[I]=-0.0119)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995	0.8040	-0.0120	0.0795	10.2642	0.8759	-0.0116	0.0807	10.9975	0.2912	-0.0122	0.0245	12.3837
	1996	0.8005	-0.0082	0.0800	10.1088	0.8686	-0.0133	0.0802	10.9963	0.2923	-0.0128	0.0255	11.9647
	1997	0.8093	-0.0125	0.0804	10.2214	0.8723	-0.0133	0.0806	10.9876	0.2885	-0.0120	0.0270	11.1296
	1998	0.8212	-0.0105	0.0822	10.1180	0.8983	-0.0122	0.0793	11.4817	0.2462	-0.0111	0.0256	10.0508
	1999	0.8114	-0.0108	0.0809	10.1632	0.8899	-0.0093	0.0789	11.3967	0.2457	-0.0114	0.0250	10.2840
	2000	0.8203	-0.0110	0.0819	10.1502	0.8971	-0.0122	0.0801	11.3521	0.2447	-0.0115	0.0269	9.5242
Space-time correlation	1998	0.8026	-0.0139	0.0802	10.1808	0.8717	-0.0118	0.0776	11.3853	0.2613	-0.0124	0.0252	10.8611
	2000	0.8228	-0.0092	0.0785	10.5987	0.8985	-0.0085	0.0822	11.0341	0.2474	-0.0106	0.0289	8.9273

Note: All statistics are significant at p=0.001.

Appendix A.5: Cluster map for average age when the highest education level was completed (AMN) in 1996, 1998 and 2000



Appendix A.6: Boxplot for standardised (Zscore) average education level completed (EMN) and average age when the highest education level completed (AMN)



Note: extreme cases and outliers are sorted by descending order: DE3, DED, DE8, DE4, DEE, DEG, DK, DE9, DEA, DE5, DE7, DEF, DEG and DEX in 1995; DE3, DED, DE8, DE4, DEE, DEG, DK, DE9, DEA, DEF, DEF, DE2, DE5, DE6 and DEX in 1996; DE3, DED, DE8, DE4, DEE, DEG, DEF, DEK, DE9, DE5, DEA, DE7, DE2 and DEX in 1997; DE3, DED, DE8, DE4, DEG, DEE, DE5, DE9, DEA, DE7, DEF, DE2, DEX, DE6, DK and DE1 in 1998; DED, DE3, DE8, DE4, DEE, DEG, DEF, DEA, DE5, DE7, DE9, DE2, DEX, DEK and DE6 in 1999; DED, DE3, DE8, DE4, DEE, DE5, DEG, DEA, DEF, DE2, DE9, DEX, DE7, DE6, DEK and DE1 in 2000 (see Appendix A.1).

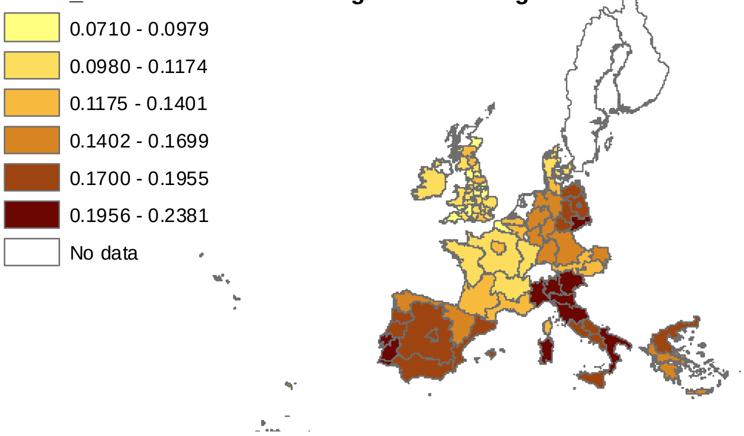
Appendix A.7: Pearson correlation between two proxies for educational attainment

	1995	1996	1997	1998	1999	2000
EMN-AMN	0.730 (0.000)** 85 0.711 (0.000)** 93	0.710 (0.000)** 85 0.695 (0.000)** 93	0.692 (0.000)** 85 0.672 (0.000)** 93	0.298 (0.000)** 85	0.269 (0.013)* 85	0.453 (0.000)** 85

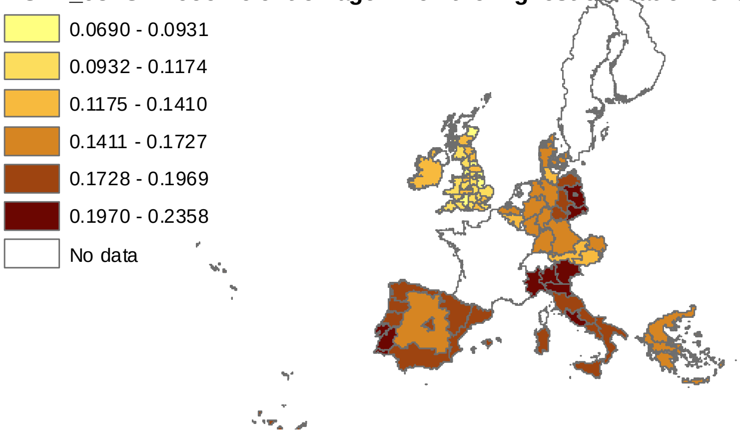
Note: ** correlation is significant at the 0.01 level (2-tailed); * correlation is significant at the 0.05 level (2-tailed).

Appendix A.8: Spatial distribution of the Gini coefficient on age when the highest education level was completed (AGINI) in 1996, 1998 and 2000

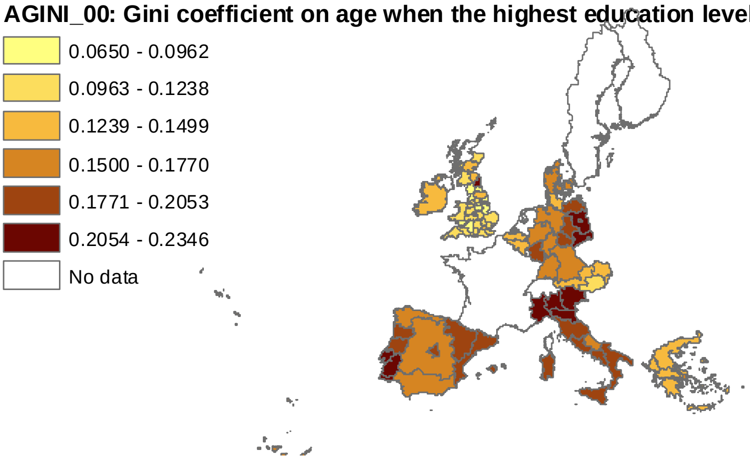
AGINI_96: Gini coefficient on age when the highest education level was completed in 1996



AGINI_98: Gini coefficient on age when the highest education level was completed in 1998

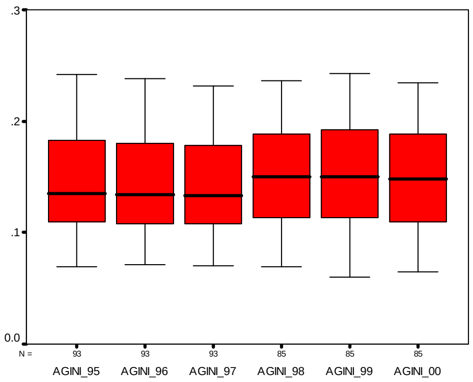


AGINI_00: Gini coefficient on age when the highest education level was completed in 2000

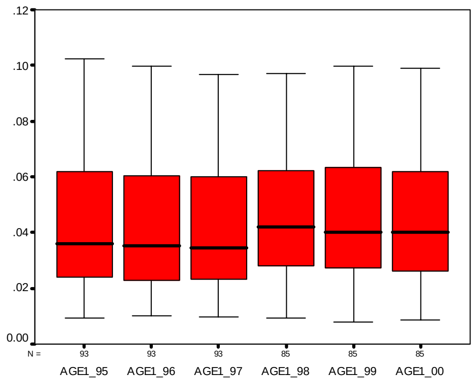


Appendix A.9: Boxplot for inequality indices on age when the highest education level was completed

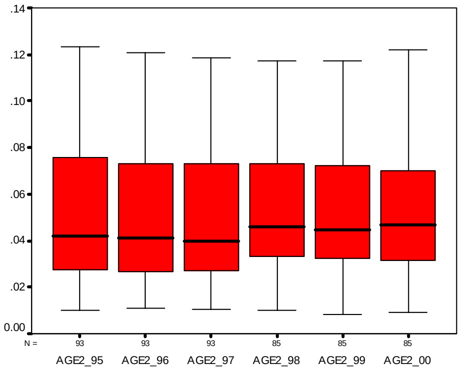
Gini coefficient (AGINI)



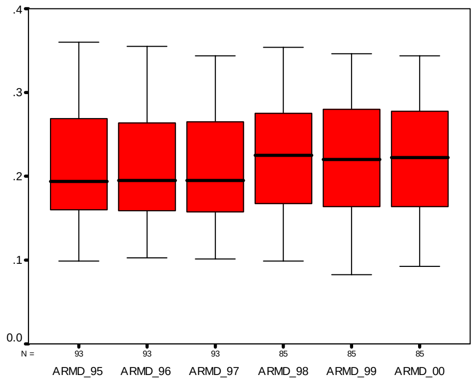
Theil index (AGE1)



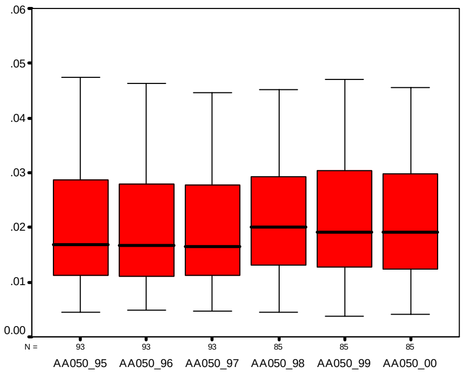
Squared coefficient of variation (AGE2)



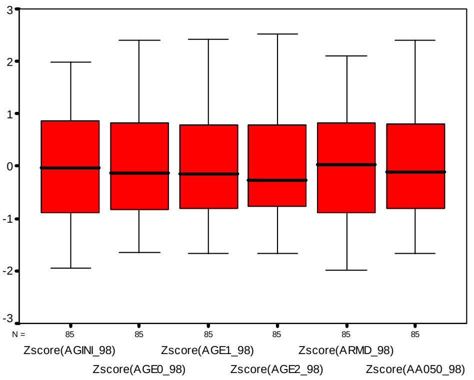
Relative mean deviation index (ARMD)



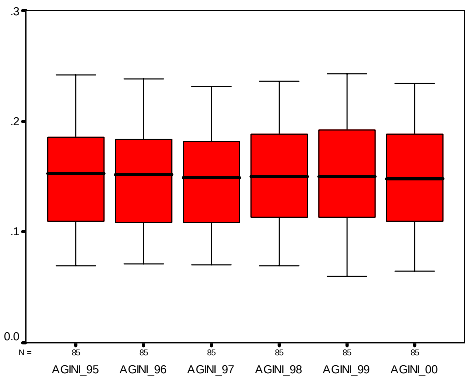
Atkinson index (AA050)



Standardized (Zscore) inequality indexes in 1998



Gini coefficient (exclude regions listwise)



Appendix A.10: The Pearson correlations among inequality indices on age when the highest grade was completed in 1998

	AGINI	AGE1	AGE2	ARMD	AA050
AGINI	1	0.980 (0.000)** 85	0.962 (0.000)** 85	0.992 (0.000)** 85	0.966 (0.000)** 85
AGE1		1	0.996 (0.000)** 85	0.966 (0.000)** 85	0.970 (0.000)** 85
AGE2			1	0.944 (0.000)** 85	0.959 (0.000)** 85
ARMD				1	0.948 (0.000)** 85
AA050					1

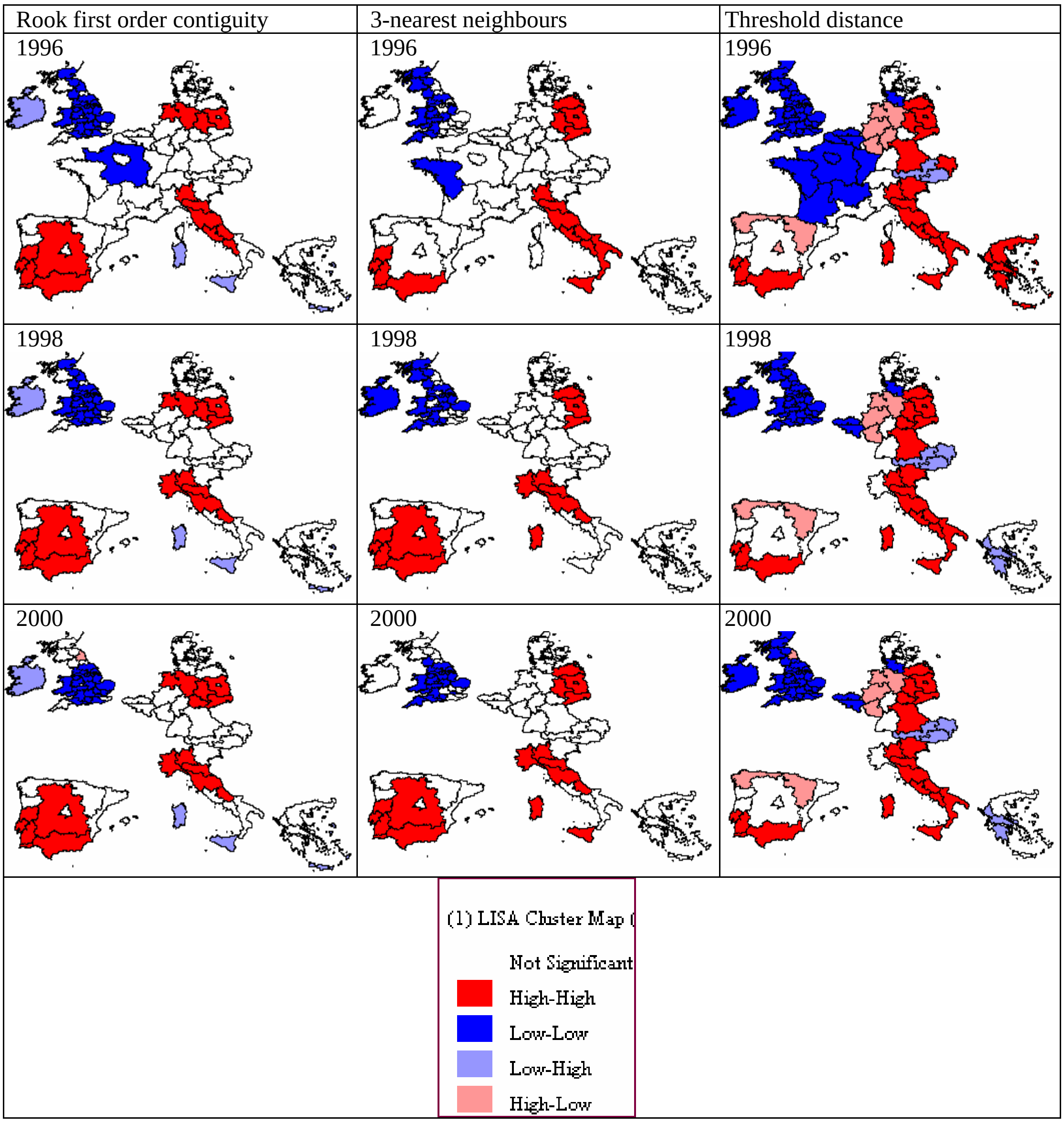
Note: ** correlation is significant at the 0.01 level (2-tailed).

Appendix A.11: Moran's I for the Gini coefficient on age when the highest education level was completed (AGINI)

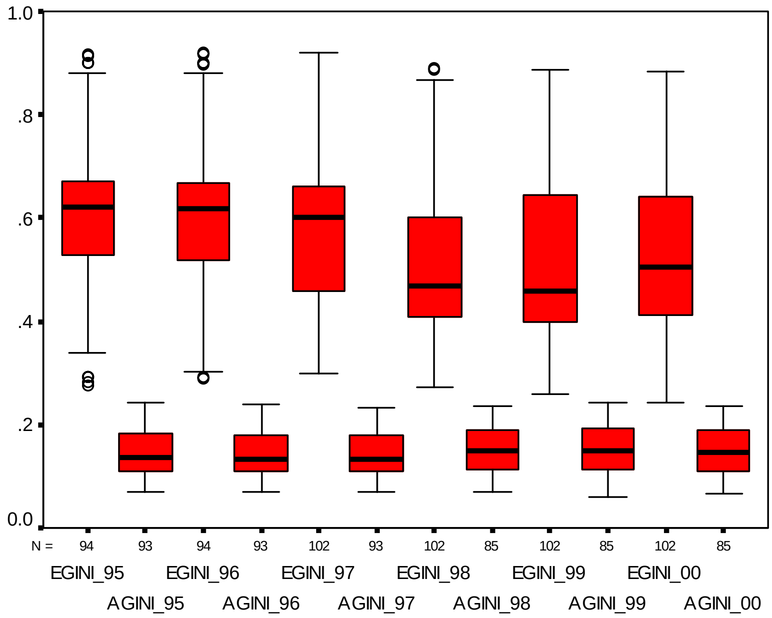
		Excluded SE LU (E[I]=-0.0109)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995	0.7366	-0.0071	0.0761	9.7727	0.7902	-0.0091	0.0797	10.0289	0.3872	-0.0115	0.0233	17.1116
	1996	0.7385	-0.0097	0.0741	10.0972	0.7913	-0.0111	0.0817	9.8213	0.3845	-0.0136	0.0214	18.6028
	1997	0.7319	-0.0079	0.0777	9.5212	0.7827	-0.0091	0.0813	9.7392	0.3807	-0.0102	0.0236	16.5636
	1998												
	1999												
	2000												
Space-time correlation	1998												
	2000												
		Excluded SE LU FR (E[I]=-0.0119)											
		rook first order contiguity				3-nearest neighbours				threshold distance			
		Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value	Moran's I	Mean	Sd	Z-value
Spatial autocorrelation	1995	0.7743	-0.0083	0.0832	9.4063	0.8516	-0.0117	0.0800	10.7913	0.4541	-0.0115	0.0269	17.3086
	1996	0.7774	-0.0121	0.0834	9.4664	0.8532	-0.0101	0.0801	10.7778	0.4519	-0.0132	0.0263	17.6844
	1997	0.7701	-0.0115	0.0805	9.7093	0.8436	-0.0161	0.0784	10.9656	0.4477	-0.0103	0.0266	17.2180
	1998	0.7557	-0.0154	0.0813	9.4846	0.8440	-0.0087	0.0799	10.6721	0.4098	-0.0132	0.0255	16.5882
	1999	0.7565	-0.0130	0.0820	9.3841	0.8393	-0.0099	0.0784	10.8316	0.4070	-0.0106	0.0261	16.0000
	2000	0.7150	-0.0108	0.0807	8.9938	0.8026	-0.0110	0.0772	10.5389	0.3527	-0.0121	0.0263	13.8707
Space-time correlation	1998	0.7696	-0.0109	0.0783	9.9681	0.8527	-0.0102	0.0795	10.8541	0.4184	-0.0113	0.0256	16.7852
	2000	0.7280	-0.0042	0.0814	8.9951	0.8089	-0.0143	0.0769	10.7048	0.3766	-0.0107	0.0256	15.1289

Note: All statistics are significant at p=0.001.

Appendix A.12: Cluster map for the Gini coefficient on age when the highest education level was completed (AGINI) in 1996, 1998 and 2000



Appendix A.13: Boxplot for Gini coefficient on education level completed (EGINI) and Gini coefficient on age when the highest education level completed (AGINI)



Note: extreme cases and outliers are sorted by descending order: EGINI: PT3, PT2, PT15 and PT14 (upper end); DE4, DEE, DED and DEG (lower end) in 1995; PT2, PT3, PT12, PT14, PT15 and PT11 (upper end); DED and DEG (lower end) in 1996; PT14 and PT12 (upper end) in 1998 (see Appendix A1.1).

Appendix A.14: Pearson correlation between two proxies for the Gini coefficient on education

	1995	1996	1997	1998	1999	2000
EGINI-AGINI	0.053 (0.627) 85 0.069 (0.510) 93	0.062 (0.576) 85 0.073 (0.486) 93	0.090 (0.412) 85 0.104 (0.319) 93	0.456 (0.000)** 85	0.443 (0.000)** 85	0.261 (0.016)* 85

Note: ** correlation is significant at the 0.01 level (2-tailed); * correlation is significant at the 0.05 level (2-tailed).