

REGIONAL ECONOMIC GROWTH AND HUMAN CAPITAL: THE ROLE OF OVEREDUCATION

Raul Ramos¹

Grup d'Anàlisi Quantitativa Regional (AQR-IREA), University of Barcelona
Dept. of Econometrics, Statistics and Spanish Economy, Avda. Diagonal 690, 08034
Barcelona (Spain). E-mail: rramos@ub.edu

Jordi Surinach

Grup d'Anàlisi Quantitativa Regional (AQR-IREA), University of Barcelona
Dept. of Econometrics, Statistics and Spanish Economy, Avda. Diagonal 690, 08034
Barcelona (Spain). E-mail: jurinach@ub.edu

Manuel Artís

Grup d'Anàlisi Quantitativa Regional (AQR-IREA), University of Barcelona
Dept. of Econometrics, Statistics and Spanish Economy, Avda. Diagonal 690, 08034
Barcelona (Spain). E-mail: manuel.artis@ub.edu

REGIONAL ECONOMIC GROWTH AND HUMAN CAPITAL: THE ROLE OF OVEREDUCATION

Abstract: This article analyses the link between human capital and regional economic growth in the European Union. Using various indicators of human capital calculated from census microdata, we conclude that the recent economic performance of European regions is associated with an increase in overeducation. In fact, measures of educational mismatch seem to be more strongly connected to regional economic performance than do other traditional measures of human capital stock.

Keywords: Regional economic growth, human capital, educational mismatch, overeducation.

JEL Classification: O18, O47, R23

CRECIMIENTO ECONÓMICO REGIONAL Y CAPITAL HUMANO: EL PAPEL DE LA SOBREENEDUCACIÓN

Resumen: El artículo analiza la relación existente entre el capital humano y el crecimiento económico regional en la Unión Europea. A partir de la consideración de varios indicadores de capital humano calculado a partir de microdatos de distintos censos, concluimos que el crecimiento económico de las regiones europeas está asociado con un incremento en la sobreeducación. De hecho, las medidas de desajuste educativo parecen estar más relacionadas con el crecimiento económico regional que otras medidas tradicionales de dotación de capital humano.

1. INTRODUCTION AND OBJECTIVES

The importance of human capital accumulation has been stressed by endogenous growth theory (Lucas, 1988 and Romer, 1990). Countries and regions with higher levels of human capital, it is argued, can expect higher growth rates than territories with lower levels. However, despite the theoretical predictions of these models, empirical evidence is inconclusive. Various explanations for this have been forwarded in the literature, but the main shortcoming seems to be that most studies have tended to rely on education as their variable for human capital, and this has usually been proxied by the average number of years of schooling or the percentage of population completing secondary or tertiary studies². In fact, recent contributions have also suggested that human capital should also consider talent, creativity and skills. For instance, Florida (2002) argues that what matters for economic growth is creativity. Creative people are workers who use their knowledge to identify problems and try to find original solutions by generating new ideas, new technology or combining both. However, there is a clear link between formal education and creativity and, in fact, most of the individuals included in the creative class are indeed more educated workers, so it is very difficult to disentangle the effect of both components of human capital (see, for instance, Marrocu and Pacci, 2012). An additional aspect in this context that has been highlighted by Rodríguez-Pose and Vilalta (2005, p. 546) is that human capital “may have little impact on regional growth if those human resources are left idle or not used to the best of their capacity in the workplace”. This argument provides an alternative explanation of the poor performance of human capital indicators in empirical growth. In particular, if the supply of highly educated workers is not matched by demand, then the impact of education on economic growth is not necessarily positive, especially if geographical labour mobility is limited. Yet, even if highly educated workers do not find a suitable job, the fact that they choose to stay in the region as unemployed or over-educated workers can represent a potentiality for economic growth. This is the central hypothesis of our research.

To date, the impact of labour market mismatch on regional economic growth has not received very much attention in the literature owing to the difficulties encountered in obtaining appropriate data to undertake such research. The only work to our knowledge that has considered this issue before is Rodríguez-Pose and Vilalta-Bufí (2005). These

authors use data from the 1994 wave of the European Community Household Panel (ECHP) to test the effects of different indicators of educational stock on regional economic growth between 1994 and 2000 for 49 regions from six EU countries (Belgium, France, Greece, Italy, Portugal, Spain and the United Kingdom). Some of these indicators are related to the matching between workers education and the labour market needs of the considered regions. In particular, they consider some qualitative and subjective indicators related to informal training, the degree of job satisfaction of workers and their perception about the infra-utilization of their skills on their current jobs. Using cross-sectional regressions, their results show that, whether it is through a more adequate use of the human resources available or through job mobility, the matching between supply and demand matters for growth.

In this paper, we take the contribution by Rodríguez-Pose and Vilalta-Bufí (2005) as a starting point and we try to extend it in several ways: first, the availability of census microdata with regional detail for a similar number of EU countries but for various time periods permits us to apply panel data techniques in order to control for unobserved regional heterogeneity, a usual criticism to cross-sectional growth regressions; second, the availability of data at different regional levels (NUTS-I, NUTS-II and NUTS-III) permits us to check the robustness of our econometric analysis to aggregation problems such as the ecological fallacy or the modifiable areal unit problem³; and, third, we will construct two different measures of overeducation based on statistical methods that overcome some of the criticism received by subjective measures of educational mismatch, such as the ones used by Rodríguez-Pose and Vilalta-Bufí (2005).

Summarising, the objective of this paper is to analyse the impact of human capital indicators, including some objective indicators of education-occupation mismatches, on regional economic growth for a wide sample of European regions using panel data techniques and considering different levels of regional aggregation.

The rest of the paper is structured as follows. First, in the section that follows, we describe the database used for the study and variable definitions. Second, the link between overeducation and regional economic growth is analysed. Finally, the paper ends with a summary of our main findings and some directions for further research.

2. MEASURING OVEREDUCATION

Educational mismatch occurs when the number of years of schooling received by a worker is higher or lower than that required to carry out his or her job. In order to measure educational mismatch, three methods employing microdata have been developed in the literature (Groot and Maassen van den Brink, 2002): the objective, the subjective and the statistical methods.

The objective method involves comparing a worker's level of education with the level required to successfully perform the functions associated with a particular post, according to a panel of experts. Workers with the same levels as those identified by this panel are classified as being "properly educated", while the rest are classified as being "mismatched".

The subjective method is based on surveys in which individuals self-classify themselves directly into either of the aforementioned categories, or alternatively surveys are conducted that enquire about the nature of workers' jobs and on the basis of these findings individuals can be classified indirectly.

The statistical method considers jobs in terms of the average number of years of education presented by workers undertaking the task and then classifies workers according to the number of years of study above or below this average plus or minus a standard deviation, or alternatively below the mode (or the corrected mode). Recently, Eurostat has also started to produce regularly a statistical indicator of educational mismatch. In particular, Eurostat (2011) defines the overqualification rate of a particular country (or region) as the percentage of employed persons in the country with a high educational level who are in low or medium skilled jobs.

Typically, the choice of one method or another tends to be determined not by theoretical arguments but rather by the availability of statistical information, as, unfortunately, these methods even when applied to the same database have been found to provide quite different results (Groot and Maassen van den Brink, 2000). Although any of the procedures has been exempted for criticism, the objective method is much more expensive and generates outdated knowledge because requirements and job titles are not

constant (Hartog, 2000) while subjective indicators seem to be particularly affected by measurement errors. In this sense, McGuinness (2006) argues that overeducated workers may be less likely to respond to questionnaires due to higher levels of job apathy which may lead to an underestimation of the incidence of overeducation. Also, workers in smaller and/or less structured organizations may lack sufficient benchmarks against which to assess their job requirements or may be applying different criteria, two factors that will lead to measurement error. Moreover, according to Hartog (2000) respondents may have a tendency to overstate the requirements of their jobs in order to upgrade the status of their position. Again, as previously mentioned, this tendency may differ between people in the same job. Summarising, although the statistical method has also some weaknesses related to the fact that it only measures realized matches, it is the most widely used in the literature (Leuven and Oosterbeck, 2011) and the one applied in this paper.

In order to conduct our research, we use microdata from the Integrated Public Use Microdata Series International (IPUMS-I) housed at the Minnesota Population Center. This provides an integrated series of census microdata samples from 1960 to the present day that has been widely used in academic works in different contexts⁴. As of January 2012, the series includes 185 samples drawn from 62 countries. 47 of these samples provide information about 13 European Union (EU) member states. Table 1 summarises the availability of information from the IPUMS-I project for these 13 EU countries. As we can see from this table, although data are available for Hungary and the Netherlands, information about the region of residence is not provided, so they are excluded from any further analysis. Moreover, for Italy and Slovenia there is only one census available, so these two countries cannot be considered in the context the panel data analysis we want to perform. Romania is not included in the analysis for a different reason: although we have more than one sample, in 1992 it has just started its transition towards democracy and a market-oriented economy making difficult the comparison with data related to 2002.

TABLE 1

One of the main advantages of using IPUMS-I data is that, when possible, variables are recoded and documented in an international context to enhance comparative research.

Educational attainment is one of these integrated variables. However, it is important to mention that there are two countries where information about educational attainment is not provided in such a comparable way. These countries are Austria and the United Kingdom. While for the case of Austria is quite straightforward to obtain a direct equivalence (although losing some detail), for the United Kingdom it is not possible, so we have had to exclude it from the analysis.

Summarising, our final sample is formed by six EU countries: Austria, France, Greece, Ireland, Portugal and Spain. Taking into account the availability of data at different regional levels (NUTS-I, NUTS-II and NUTS-III) and in order to avoid the problems derived from aggregating data or choosing a particular level of regional detail and to check the robustness of our results, we work with three different samples: 26 NUTS-I regions, 72 NUTS-II regions and 164 NUTS-III regions, which are shown in the maps shown in annex 1. The last columns in table 1 show the size of the country samples during the two most recent censuses conducted in European countries: 1990-1991-1992 and 1999-2001-2002. The total number of individuals considered in the analysis totals more than 6 million.

As previously mentioned The obvious advantage of using the IPUMS-I samples as opposed to those supplied directly by the National Institute of Statistics lies in the fact that a number of key variables such as educational level⁵ and occupations⁶ are recoded using a homogenous classification. Drawing on this information, it is possible to calculate statistical measures of educational mismatch at the individual level and, subsequently, to obtain regional indicators of the incidence and intensity of overeducation.

The first step in calculating a measure of educational mismatch involves transforming educational levels into the corresponding number of years of schooling⁷. The average number of schooling years for the working population in the countries and time periods considered are quite similar to those obtained elsewhere (see, for example, Barro and Lee, 2000): the number of schooling years increased substantially between the eighties and nineties in all European countries, but the greatest increase was recorded in countries with the lowest initial levels. We also drew on the census information to calculate the percentage number of workers that had completed secondary and tertiary

studies. An analysis of these indicators shows that the increase in the number of schooling years reflected a higher rate of enrolment at both levels of education. Here again, our findings are similar to those when drawing on information contained in other databases, including the World Bank's World Development Indicators or Eurostat Eurostat's Regional Statistics Census data⁸.

Having transformed the information regarding levels of education attained into number of schooling years, we then compare an individual's number of schooling years with the number required to undertake his or her job. Specifically, individual i working in occupation j and living in region y of country z at time t is considered to be properly educated if his number of years of schooling is equal to the most frequent value (the mode) for the number of schooling years for workers in occupation j in sector k of country z ⁹. If the number of schooling years is higher/lower than the mode, then the individual is classified as being over-/under-educated. Information at the regional level is obtained by aggregating the individual data.

A somewhat different perspective on this educational mismatch is obtained if we focus on the intensity of under- and overeducation rather than on its incidence. This is achieved by breaking the number of schooling years down into three components: the number of years of overeducation, the number of years required for a particular post and the number of years of undereducation¹⁰. These two groups of indicators constitutes our first statistical measures of regional overeducation to be used in the empirical analysis¹¹.

As previously mentioned, the second measure is based on Eurostat's methodology. In particular, the overqualification rate of a particular region has been calculated as the percentage of employed persons in that region with a high educational level (having completed tertiary education, ISCED 5 or 6) who are in low or medium skilled jobs (ISCO occupation levels 4 to 9)¹².

Table 2 provides summary statistics for the different human capita indicators described above. As we can see from this table, the proportion of workers with secondary and tertiary studies and the number of schooling year have clearly increased during the considered period and regional differences (both in absolute and relative terms) have decreased. However, this increase in traditional human capital indicators has also been

accompanied by an upward trend on educational mismatch. We can see that, both, the proportion of overqualified and overeducated workers have increased and that the intensity of overeducation (measured as the “excess of schooling years”) is also higher when using more recent data¹³.

TABLE 2

Summarising, the simple descriptive analysis conducted in this section shows that there has been a marked increase in levels of schooling in the EU regions and that this has reduced regional differences in human capital stock. However, this increase in human capital has not been accompanied by a similar increase in the number of qualified jobs, and, as a result, both the incidence and intensity of overeducation have increased across EU regions. The next section analyses the effects of these two complementary trends on regional economic growth.

3. OVEREDUCATION AND REGIONAL ECONOMIC GROWTH

This section examines the link between educational mismatch and regional economic growth in the EU. In order to disentangle the effect of this mismatch on growth, we first consider the effects of traditional indicators of human capital stock (number of schooling years, percentage of workers with secondary and tertiary studies) and, then, turn our attention to the effects of the different indicators related to mismatches between education and occupation.

In order to determine which human capital measures have the greatest impact on regional economic growth and to analyse the effects of educational mismatch, we estimated panel data models using the variation between 1995, 2000 and 2005 in Gross Domestic Product (GDP) per capita data adjusted for Purchasing Power Parities (PPP) provided by Eurostat as the endogenous variable. It should be borne in mind that the main advantage of adopting a panel data approach is that it allows us to control for unobservable heterogeneity through the inclusion of regional and time fixed effects.

Specifically, GDP per capita growth between 1995 and 2000 and between 2000 and 2005 is regressed on the initial level of GDP per capita and the human capital indicators

for the different sets of regions for which both GDP and human capital data are available for both periods. In particular:

$$\left(\ln y_{i,t} - \ln y_{i,t-\tau}\right) / x = \alpha + \beta \cdot \ln y_{i,t-\tau} + \gamma \cdot x_{i,t-\tau} + \eta_t + \mu_i + \varepsilon_{i,t} \quad (1)$$

where $\ln y_{i,t}$ is the logarithm of GDP in region i at time t , $x_{i,t-\tau}$ represents the different human capital indicators at time $t-\tau$ ¹⁴, η_t is a time fixed effect, μ_i a region fixed effect, and $\varepsilon_{i,t}$ a random error term that varies across regions and time periods.

Table 3 summarises the results of estimating equation (1) for the different sets of regions (NUTS-I, NUTS-II and NUTS-III) included in our analysis. The different columns of the table show the results obtained when estimating models with the different explanatory variables. In particular, in models 1 and 2, growth was regressed on initial GDP per capita and traditional human capital indicators: the percentage of working population with secondary and tertiary studies and the number of schooling years. Indicators of educational mismatch are included in models 3, 4 and 5. Model 3 includes the number of schooling years together with the overqualification rate calculated following the Eurostat's methodology. The percentage of properly educated workers and the percentage of over-educated workers calculated using the statistical method are included in model 4, while in model 5 the number of schooling years is broken down in terms of required, over and under.

TABLE 3

The results in Table 3 reveal a number of interesting results. First, the coefficient of initial GDP per capita is always negative and significant at the usual levels, indicating that a process of regional convergence has occurred during the period under review. This process is still apparent when the various human capital indicators are included.

The introduction of the traditional indicators of human capital in models 1 and 2 reveals their positive impact on economic growth. The coefficients are always positive and statistically significant at the 1% level for secondary studies and schooling years, but not for tertiary studies. For the later, evidence is less clear, particularly when working

with NUTS-I regions. Probably, this result is related with the greater incidence of overeducation among graduates.

In model 3, the percentage of overqualified workers is included in the specification together with the average number of schooling years. For both variables, the two coefficients are positive and statistically significant at the 1% level, with the only exception of the NUTS-III sample where the proportion of overqualified workers is only significant at the 10% level.

In model 4, the percentage of properly educated workers and the percentage of over-educated workers are included in the regression. The coefficients associated to the percentage of properly educated workers are positive and statistically significant at the 10% in all cases, a result which is in line with previous evidence. However, the evidence for the proportion of over-educated workers is less clear depending of the considered sample. While it is not statistically significant at the usual levels for the NUTS-I and NUTS-II samples, the results for the NUTS-III show a positive and significant effect at the 5% level. It is worth mentioning that, in this case, the magnitude of the coefficient associated with the percentage of over-educated workers is greater than that associated with the percentage of properly educated workers. This result lends some support to the hypothesis that at the regional level (albeit not necessarily at the individual level) overeducation might be seen more as an investment than as a cost¹⁵. This evidence is much clearer when looking at model 5.

Model 5 show the results of breaking down the number of schooling years into required years of schooling, years of infraeducation and years of overeducation. The results of this model confirm previous evidence but also provide an additional interesting finding: there is a positive and significant effect of the average number of required years and the average number of years of overeducation, while the average number of years of undereducation has a negative and significant effect.

Several checks have been carried out in order to guarantee the robustness of the results. First, table 4 shows the results of estimating equation (1) using a common sample of countries. The main difference with previous results is that now French regions are not included in the analysis for the NUTS-I and NUTS-II samples. As we can see from the

table, the results are in line with the ones explained above. A second robustness check has consisted in estimating the model without Austrian regions as in this case, there was not a perfect equivalence between Austrian educational levels and the ones used by IPUMS-I to standardize the variables. The results in table 5 are again similar to the ones detailed above.

TABLES 4 AND 5

Two additional robustness checks have been carried out in order to test the stability of the results to changes in the econometric specification: first, we have substituted regional fixed effects by regional random effects and, second, we have added some additional regional time-varying control variables that have been calculated from the IPUMS-I census samples. In particular, the additional control variables include regional demographic structure (percentage of population over the age of 50), labour market characteristics (inactivity rate) and employment composition (agriculture, manufacturing, building)^{16, 17}. Results are shown, respectively, in tables 6 and 7. Again the conclusions are similar.

TABLES 6 AND 7

4. FINAL REMARKS

While the limited time frame and the nature of the analysis mean that any conclusions drawn here should be considered with caution, the study does seem to indicate the presence of a significant correlation between overeducation and regional economic performance in recent years. The impact of overeducation on an individual's earnings is well known: he will tend to earn less than his "properly educated" counterparts. However, at the regional level, our results indicate a more favourable picture: overeducated workers represent an opportunity to take advantage of the generation of more qualified jobs. This finding does not differ greatly from those reported in studies analysing the differences between private and social returns to schooling (see, for example, Moretti, 2004). In a recent study comparing various EU countries, Middendorf (2008) also found that returns to schooling are significantly and negatively related to the

educational attainment of the population, a result which is in line with the findings reported herein.

From a policy perspective, our results indicate that even when qualified workers are unable to find a suitable job, they are still more productive at the aggregate level than their unqualified counterparts. This implies that there is a good case for public investment in education, even though a number of recent studies fail to provide favourable evidence regarding the link between human capital and growth. However, in a context of high geographical mobility, regions will not benefit directly from their “over-investment” in the education of their population. In this sense, one aspect that has not been considered in this paper is the probable existence of spatial spillovers of human capital (Tselios, 2008; Olejnick, 2008). This certainly constitutes a potential line for future research and one that needs to be considered from a policy perspective. Finally, we should stress (as Rodriguez-Pose and Vilalta-Bufí, 2005, have done so before) that the use of microeconomic data in constructing regional indicators of educational mismatch represents a step forward with respect to the traditional indicators of human capital, but in this area a considerable amount of work has still to be done.

5. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 216813. The authors gratefully acknowledge the support received from the Spanish Ministry of Science through the projects ECO2010-16006 and ECO2009-12678. The usual disclaimer applies.

This research is based on the IPUMS-International database (Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis: University of Minnesota, 2011). The authors wish to thank the following statistics offices for providing the underlying data that made this research possible: the National Bureau of Statistics, Austria; the National Institute of Statistics and Economic Studies, France; the National Statistics Office, Greece; the Central Statistics Office, Ireland; the National Institute of Statistics, Portugal; and, the National Institute of Statistics, Spain.

6. REFERENCES

Barro, R., Lee, J.W. (2000), *International data on educational attainment: Updates and implications*, CID Working Paper 42, Harvard University.

Croce, G., Ghignoni, E. (2011), “Overeducation and spatial flexibility in Italian local labour markets”, *MPRA Paper 29670*, University Library of Munich, Germany.

Büchel, F., Van Ham, M. (2003), “Overeducation, Regional labor markets, and spatial flexibility”, *Journal of Urban Economics*, vol. 53, pp. 482-493.

De la Fuente, A., Doménech, R. (2006), “Human capital in growth regressions: How much difference does data quality make?”, *Journal of the European Economic Association*, vol 4, pp. 1-36.

Duncan, G. J., Hoffmann, S. D. (1981), “The incidence and wage effects of overeducation”, *Economics of Education Review*, vol 1, pp. 75-86.

Duque, J. C., Artís, M., Ramos, R. (2006), “The ecological fallacy in a time series context: evidence from Spanish regional unemployment rates”, *Journal of Geographical Systems*, vol. 8, pp. 391-410.

Eurostat (2011), *Migrants in Europe - A statistical portrait of the first and second generation*. DOI 10.2785/5318.

Florida R. (2002), *The rise of the creative class and how it's transforming work, leisure, community, and everyday life*, New York: Basic Books.

Frank, R. (1978), “Why Women Earn Less: The Theory and Estimation of Differential Overqualification”, *American Economic Review*, vol. 68(3), pp. 360-373.

Groot, W. and Maassen Van den Brink, H. (2000), “Overeducation in the Labor Market: A Meta-Analysis”, *Economics of Education Review*, vol. 19, pp. 149-158.

Groot, W. and Maassen Van Den Brink, H. (2002), “Skill Mismatches in the Dutch Labor Market”, *International Journal of Manpower*, vol. 21 (8), pp. 584-595.

Hartog, J. (2000), “Overeducation and earnings: where are we, where should we go?”, *Economics of Education Review*, vol. 19, pp. 131-147.

Hensen, M., de Vries, M. R., Cörvers, F. (2009), “The role of geographic mobility in reducing education-job mismatches in the Netherlands”, *Papers in Regional Science*, vol. 88 (3), pp. 667-682.

Jauhiainen, S. (2011), “Overeducation in the Finnish regional labour markets”, *Papers in Regional Science*, vol. 90 (3), pp. 573–588.

Lenton, P. (2012), “Over-education across British Regions”, *Regional Studies*, forthcoming.

Leuven, E., Oosterbeek, H. (2011), Overeducation and Mismatch in the Labor Market in Hanushek, E. and Welch, F. (eds.): *Handbook of the Economics of Education*, 4, 283-326, Elsevier Science.

Lucas, R.E. (1988), “On the mechanics of economic development”, *Journal of Monetary Economics*, vol. 22, pp. 3-42.

Marroccu, E., Pacci, R. (2012), “Education or Creativity: what matters most for economic performance?”, *Journal of Economic Geography*, forthcoming.

McGuinness, S. (2006), “Overeducation in the labour market”, *Journal of Economic Surveys*, 20(3), pp. 387-418.

Middendorf, T. (2008), *Returns to Education in Europe*, Ruhr Economic Papers, #65.

Minnesota Population Center (2011), Integrated Public Use Microdata Series, International: Version 6.1 [Machine-readable database]. Minneapolis: University of Minnesota.

Moretti, E. (2004), “Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data”, *Journal of Econometrics*, vol. 121 (1-2), pp. 175-212.

Olejnick, A. (2008), “Using the spatial autoregressively distributed lag model in assessing the regional convergence of per-capita income in the EU25”, *Papers in Regional Science*, vol. 87 (3), pp. 371-385.

Quinn, M. A., Rubb, S. (2011), “Spouse overeducation and family migration: evidence from the US”, *Journal of Family and Economic Issues*, vol. 32 (1), pp. 36-45.

Rodríguez-Pose, A., Vilalta-Buñi, M. (2005), “Education, migration, and job satisfaction: the regional returns of human capital in the EU”, *Journal of Economic Geography*, vol. 5, pp. 545–566.

Romer, P. (1989), *Human capital and growth: Theory and evidence*, NBER Working Paper 3173.

Sanromá, E., Ramos, R. (2011), “Overeducation and local labour markets in Spain”, IZA DP 6028.

Temple, J. (2001), “Growth effects of education and social capital in the OECD countries”, *OECD Economic Studies*, vol. 33, pp. 57-101.

Tselios, V. (2008), “Income and educational inequalities in the regions of the European Union: Geographical spillovers under welfare state restrictions”, *Papers in Regional Science*, vol. 87 (3), pp. 403-431.

7. TABLES

Table 1. Availability of microdata samples for EU countries from the IPUMS-I project

Country	Years	Number of regions			Included	Number of individuals	
		NUTS I	NUTS II	NUTS III		1990-1991-1992	1999-2001-2002
Austria	1971, 1981, 1991, 2001	3	9	31	X	345,004	370,179
France	1962, 1968, 1975, 1982, 1990, 1999, 2006	8	22	---	X	932,384	1,156,454
Germany (East / West)	1970, 1971, 1981, 1987	11	---	---		---	---
Greece	1971, 1981, 1991, 2001	4	13	51	X	327,529	381,334
Hungary	1970, 1980, 1990, 2001	---	---	---		---	---
Ireland	1971, 1979, 1981, 1986, 1991, 1996, 2002, 2006	1	2	8	X	353,149	410,688
Italy	2001	5	19	---		---	---
Netherlands	1960, 1971, 2001	---	---	---		---	---
Portugal	1981, 1991, 2001	3	7	22	X	199,685	227,712
Romania	1977, 1992, 2002	4	8	---		---	---
Slovenia	2002	1	2	12		---	---
Spain	1981, 1991, 2001	7	19	52	X	626,202	742,777
United Kingdom	1991, 2001	12	---	---			

Source: Authors' own calculations based on IPUMS-I microdata (January 2012)

<https://international.ipums.org/international/>

Table 2. Descriptive statistics

NUTS I (26 regions)	Year	Mean	Std. Dev.	Min	Max
Proportion of workers with secondary studies	1991	0.157	0.106	0.049	0.419
	2001	0.210	0.115	0.075	0.464
Proportion of workers with tertiary studies	1991	0.059	0.028	0.013	0.145
	2001	0.091	0.044	0.042	0.214
Schooling years	1991	5.848	1.203	3.479	8.126
	2001	7.228	1.186	4.537	9.692
Proportion of overqualified workers	1991	0.324	0.139	0.024	0.554
	2001	0.365	0.149	0.043	0.608
Proportion of overeducated workers	1991	0.332	0.112	0.207	0.559
	2011	0.392	0.093	0.204	0.580
Years of overeducation	1991	2.066	0.927	0.956	4.035
	2001	2.850	0.989	1.085	4.418
<hr/>					
NUTS II (72 regions)	Year	Mean	Std. Dev.	Min	Max
Proportion of workers with secondary studies	1991	0.154	0.102	0.049	0.437
	2001	0.209	0.111	0.075	0.474
Proportion of workers with tertiary studies	1991	0.054	0.022	0.013	0.145
	2001	0.083	0.037	0.032	0.214
Schooling years	1991	5.789	1.093	3.479	8.431
	2001	7.158	1.053	4.537	9.692
Proportion of overqualified workers	1991	0.338	0.142	0.023	0.613
	2001	0.376	0.158	0.039	0.637
Proportion of overeducated workers	1991	0.323	0.121	0.164	0.578
	2011	0.384	0.102	0.180	0.603
Years of overeducation	1991	2.012	0.974	0.805	4.257
	2001	2.768	1.046	0.937	4.692
<hr/>					
NUTS III (164 regions)	Year	Mean	Std. Dev.	Min	Max
Proportion of workers with secondary studies	1991	0.179	0.114	0.027	0.439
	2001	0.250	0.109	0.049	0.486
Proportion of workers with tertiary studies	1991	0.040	0.016	0.007	0.095
	2001	0.062	0.025	0.020	0.172
Schooling years	1991	5.836	1.256	3.073	8.431
	2001	7.191	1.190	3.945	9.692
Proportion of overqualified workers	1991	0.345	0.180	0.011	0.642
	2001	0.405	0.200	0.030	0.698
Proportion of overeducated workers	1991	0.321	0.138	0.118	0.597
	2011	0.363	0.122	0.145	0.616
Years of overeducation	1991	1.899	1.116	0.584	4.307
	2001	2.319	1.114	0.762	4.926

Source: Authors' own calculations based on IPUMS-I microdata (January 2012)

Table 3. Estimation results of panel data models with spatial and time fixed effects

NUTS I – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.229***	-0.217***	-0.218***	-0.232***	-0.210***
% workers with secondary studies	0.414***	-	-	-	-
% workers with tertiary studies	0.320*	-	-	-	-
Schooling years	-	0.025***	0.027***	-	-
% of overqualified workers	-	-	0.081***	-	-
% of adequately educated workers	-	-	-	0.126*	-
% of overeducated workers	-	-	-	0.069	-
Years of infraeducation	-	-	-	-	-0.052***
Required years of schooling	-	-	-	-	0.037***
Years of overeducation	-	-	-	-	0.023***
R squared	0.931	0.914	0.934	0.924	0.951
Observations	52	52	52	52	52
NUTS II – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.239***	-0.220***	-0.218***	-0.218***	-0.228***
% workers with secondary studies	0.390***	-	-	-	-
% workers with tertiary studies	0.194*	-	-	-	-
Schooling years	-	0.020***	0.019***	-	-
% of overqualified workers	-	-	0.033*	-	-
% of adequately educated workers	-	-	-	0.064*	-
% of overeducated workers	-	-	-	0.038	-
Years of infraeducation	-	-	-	-	-0.035***
Required years of schooling	-	-	-	-	0.030***
Years of overeducation	-	-	-	-	0.017***
R squared	0.907	0.862	0.867	0.834	0.901
Observations	144	144	144	144	144
NUTS III – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.268***	-0.280***	-0.280***	-0.287***	-0.280***
% workers with secondary studies	0.405***	-	-	-	-
% workers with tertiary studies	0.208**	-	-	-	-
Schooling years	-	0.029***	0.029***	-	-
% of overqualified workers	-	-	0.001*	-	-
% of adequately educated workers	-	-	-	0.067*	-
% of overeducated workers	-	-	-	0.143**	-
Years of infraeducation	-	-	-	-	-0.031***
Required years of schooling	-	-	-	-	0.026***
Years of overeducation	-	-	-	-	0.030***
R squared	0.822	0.845	0.845	0.845	0.848
Observations	328	328	328	328	328

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4. Estimation results of panel data models with spatial and time fixed effects
with a common sample of countries

NUTS I – all regions excluding France					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.239***	-0.256***	-0.251***	-0.270***	-0.249***
% workers with secondary studies	0.335*	-	-	-	-
% workers with tertiary studies	0.330*	-	-	-	-
Schooling years	-	0.028***	0.029***	-	-
% of overqualified workers	-	-	0.048*	-	-
% of adequately educated workers	-	-	-	0.100*	-
% of overeducated workers	-	-	-	0.122	-
Years of infraeducation	-	-	-	-	-0.035*
Required years of schooling	-	-	-	-	0.027***
Years of overeducation	-	-	-	-	0.029***
R squared	0.923	0.955	0.962	0.960	0.956
Observations	36	36	36	36	36
NUTS II – all regions excluding France					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.241***	-0.258***	-0.259***	-0.269***	-0.268***
% workers with secondary studies	0.347***	-	-	-	-
% workers with tertiary studies	0.181*	-	-	-	-
Schooling years	-	0.024***	0.024***	-	-
% of overqualified workers	-	-	0.010*	-	-
% of adequately educated workers	-	-	-	0.120**	-
% of overeducated workers	-	-	-	0.101**	-
Years of infraeducation	-	-	-	-	-0.017
Required years of schooling	-	-	-	-	0.019***
Years of overeducation	-	-	-	-	0.027***
R squared	0.883	0.901	0.902	0.899	0.905
Observations	100	100	100	100	100

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Estimation results of panel data models with spatial and time fixed effects excluding Austrian regions

NUTS I – all regions excluding Austria					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.237***	-0.207***	-0.197***	-0.241***	-0.209***
% workers with secondary studies	0.382***	-	-	-	-
% workers with tertiary studies	0.265*	-	-	-	-
Schooling years	-	0.037**	0.054***	-	-
% of overqualified workers	-	-	0.120***	-	-
% of adequately educated workers	-	-	-	0.052*	-
% of overeducated workers	-	-	-	0.081	-
Years of infraeducation	-	-	-	-	-0.052***
Required years of schooling	-	-	-	-	0.037***
Years of overeducation	-	-	-	-	0.024*
R squared	0.931	0.914	0.952	0.895	0.948
Observations	46	46	46	46	46
NUTS II – all regions excluding Austria					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.245***	-0.219***	-0.217***	-0.219***	-0.224***
% workers with secondary studies	0.359***	-	-	-	-
% workers with tertiary studies	0.142*	-	-	-	-
Schooling years	-	0.021**	0.022*	-	-
% of overqualified workers	-	-	0.035*	-	-
% of adequately educated workers	-	-	-	0.060*	-
% of overeducated workers	-	-	-	0.031	-
Years of infraeducation	-	-	-	-	-0.034***
Required years of schooling	-	-	-	-	0.022**
Years of overeducation	-	-	-	-	0.007
R squared	0.905	0.849	0.855	0.840	0.896
Observations	126	126	126	126	126
NUTS III – all regions excluding Austria					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.277***	-0.282***	-0.283***	-0.283***	-0.280***
% workers with secondary studies	0.272***	-	-	-	-
% workers with tertiary studies	0.134*	-	-	-	-
Schooling years	-	0.026***	0.025***	-	-
% of overqualified workers	-	-	0.007*	-	-
% of adequately educated workers	-	-	-	0.014*	-
% of overeducated workers	-	-	-	0.134***	-
Years of infraeducation	-	-	-	-	-0.029***
Required years of schooling	-	-	-	-	0.020***
Years of overeducation	-	-	-	-	0.024***
R squared	0.826	0.838	0.838	0.835	0.841
Observations	266	266	266	266	266

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6. Estimation results of panel data models with country
and time fixed effects and regional random effects

NUTS I – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.093***	-0.084***	-0.094***	-0.016	-0.102***
% workers with secondary studies	0.282***	-	-	-	-
% workers with tertiary studies	0.266***	-	-	-	-
Schooling years	-	0.020***	0.023***	-	-
% of overqualified workers	-	-	0.040*	-	-
% of adequately educated workers	-	-	-	0.155*	-
% of overeducated workers	-	-	-	0.140*	-
Years of infraeducation	-	-	-	-	-0.057***
Required years of schooling	-	-	-	-	0.030***
Years of overeducation	-	-	-	-	0.015**
Observations	52	52	52	52	52
NUTS II – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.065***	-0.051***	-0.050***	-0.023**	-0.058***
% workers with secondary studies	0.229***	-	-	-	-
% workers with tertiary studies	0.156**	-	-	-	-
Schooling years	-	0.011***	0.011***	-	-
% of overqualified workers	-	-	0.003*	-	-
% of adequately educated workers	-	-	-	0.089*	-
% of overeducated workers	-	-	-	0.094*	-
Years of infraeducation	-	-	-	-	-0.034***
Required years of schooling	-	-	-	-	0.017***
Years of overeducation	-	-	-	-	0.006**
Observations	144	144	144	144	144
NUTS III – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.071***	-0.081***	-0.079***	-0.055***	-0.076***
% workers with secondary studies	0.268***	-	-	-	-
% workers with tertiary studies	0.151**	-	-	-	-
Schooling years	-	0.019***	0.018***	-	-
% of overqualified workers	-	-	0.016*	-	-
% of adequately educated workers	-	-	-	0.133**	-
% of overeducated workers	-	-	-	0.206***	-
Years of infraeducation	-	-	-	-	-0.047***
Required years of schooling	-	-	-	-	0.025***
Years of overeducation	-	-	-	-	0.007*
Observations	328	328	328	328	328

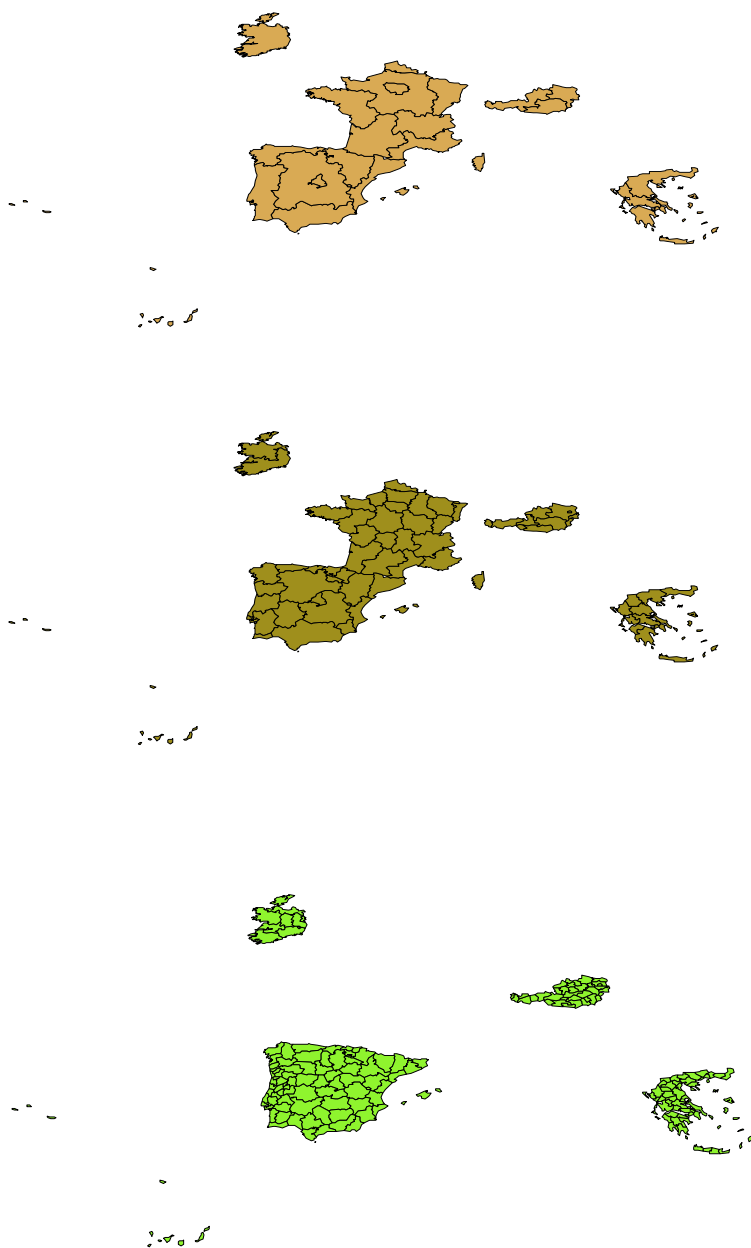
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7. Estimation results of panel data models with regional time varying controls and regional random effects

NUTS I – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.130***	-0.131***	-0.122***	-0.086***	-0.112***
% workers with secondary studies	0.139***	-	-	-	-
% workers with tertiary studies	0.066*	-	-	-	-
Schooling years	-	0.010**	0.009**	-	-
% of overqualified workers	-	-	0.001*	-	-
% of adequately educated workers	-	-	-	0.017*	-
% of overeducated workers	-	-	-	0.003	-
Years of infraeducation	-	-	-	-	-0.036**
Required years of schooling	-	-	-	-	0.010**
Years of overeducation	-	-	-	-	0.001*
Observations	52	52	52	52	52
NUTS II – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.099***	-0.097***	-0.091***	-0.063***	-0.088***
% workers with secondary studies	0.082***	-	-	-	-
% workers with tertiary studies	0.004*	-	-	-	-
Schooling years	-	0.005**	0.004*	-	-
% of overqualified workers	-	-	0.003*	-	-
% of adequately educated workers	-	-	-	0.031*	-
% of overeducated workers	-	-	-	0.003	-
Years of infraeducation	-	-	-	-	-0.024**
Required years of schooling	-	-	-	-	0.006**
Years of overeducation	-	-	-	-	0.003
Observations	144	144	144	144	144
NUTS III – all regions					
Annualised GDPpc growth	(1)	(2)	(3)	(4)	(5)
Initial GDPpc	-0.073***	-0.092***	-0.084***	-0.050***	-0.087***
% workers with secondary studies	0.026*	-	-	-	-
% workers with tertiary studies	0.145**	-	-	-	-
Schooling years	-	0.007***	0.005***	-	-
% of overqualified workers	-	-	0.027**	-	-
% of adequately educated workers	-	-	-	0.081*	-
% of overeducated workers	-	-	-	0.054***	-
Years of infraeducation	-	-	-	-	-0.027***
Required years of schooling	-	-	-	-	0.009***
Years of overeducation	-	-	-	-	0.006*
Observations	328	328	328	328	328

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Annex 1. Countries and regions included in the analysis
(NUTS-I, NUTS-II, NUTS-III)



Source: Authors' elaboration from Eurostat

http://epp.eurostat.ec.europa.eu/portal/page/portal/gisco_Geographical_information_maps/popups/references/administrative_units_statistical_units_1

Annex 2. Equivalence between educational levels and number of schooling years

<u>Educational levels</u>	<u>Number of schooling years</u>
Less than primary completed (n.s.)	0
No schooling	0
Some primary completed	3
Primary (4 yrs) completed	4
Primary (5 yrs) completed	5
Primary (6 yrs) completed	6
Lower secondary general completed	8
Lower secondary technical completed	10
Secondary, general track completed	12
Some college completed	13
Secondary or post-secondary technical completed	13
Secondary, technical track completed	15
Post-secondary technical education	16
University completed	17

Source: Authors' elaboration from IPUMS-I (*edattand* variable)

<https://international.ipums.org/international/>

END NOTES

¹ Corresponding author.

² The quality of these data has also been called into question (De la Fuente and Doménech, 2006).

³ See Duque et al. (2006) for a review.

⁴ <http://bibliography.ipums.org>

⁵ Easily obtained from the nine homogenous categories that are considered: Less than primary completed / Some primary completed / Primary (6 yrs) completed / Lower secondary general completed / Secondary, general track completed / Some college completed / Secondary, technical track completed / Post-secondary technical education / University completed.

⁶ At a 3-digit level of detail (more than 400).

⁷ As previously mentioned, as schooling levels in each country have been homogenised as part of the IPUMS-I project, the equivalence between educational levels and number of schooling years is quite straightforward and it is shown in annex 2.

⁸ However, it is worth mentioning that there are significant differences between our educational indicators and Eurostat's Regional Education Statistics. The results of the different comparisons mentioned in the text are available from the authors on request.

⁹ Note, we assume that the educational requirements of a certain workplace are identical across regions in the same country, but that they can vary over time. A similar assumption is adopted by the few studies analysing overeducation at the regional level such as Jauhiaine (2011) and Lenton (2012).

¹⁰ In a similar way to the ORU specification of the Mincer equation commonly used in the economics of education literature (Duncan and Hoffman, 1981).

¹¹ Studies considering the regional dimension of overeducation are scarce. In fact, the relationship between overeducation and geography has been systematically ignored in the literature. To our knowledge, the only link between the two is related to the differential overqualification theory which suggests that overeducation basically affects married women since their job search is restricted to the local labour market in which they live, while the husband is able to search for a job more in keeping with his level of education in a wider labour market (Frank, 1978). Recent studies from this perspective include Büchel and Van Ham (2003) for Germany, Hensen et al. (2009) for the Netherlands, Quinn and Rubb (2011) for the United States, Sanromá and Ramos (2011) for Spain and Croce and Ghignoni (2011) for Italy. It is worth mentioning that, although these studies take into account the geographical dimension in their analysis, all of them focus on the individual level.

¹² <http://www.ilo.org/public/english/bureau/stat/isco/isco88/major.htm>

These groups include the following:

4 Clerks

5 Service workers and shop and market sales workers

6 Skilled agricultural and fishery workers

7 Craft and related trades workers

8 Plant and machine operators and assemblers

9 Elementary occupations

¹³ Full details of the results at the regional level are available from the authors on request.

¹⁴ As Temple (2001) highlights, this specification is preferred to the analysis of the relation between the change in output and the change in education as in this case causality could run from output (or anticipated output) to education, and not vice versa. As long-run changes in average educational attainment are driven by government policy, it seems plausible that as output and tax revenues increase, governments will often allocate more resources to education, and attainment will rise for a transitional period. This critique does not apply to the specification between output growth and the initial level of human capital as considered here. The use of schooling years (instead of enrolment rates) (and the use of panel data) makes it more unlikely that reverse causation could explain the positive and significant effects of human capital and growth (de la Fuente and Domenech, 2006).

¹⁵ This result is robust to the inclusion of the average schooling levels in the region as an additional control variable. The reason for including this control is that it might be thought that the positive and significant sign of the percentage of overeducated workers could be related to the greater presence of educated workers.

¹⁶ It would have been interesting to have included controls related to regional innovation capacity, but information from Eurostat is not available for the considered regions.

¹⁷ Full results are available from the authors on request.