HUMAN CAPITAL ACCUMULATION AND GEOGRAPHY: EMPIRICAL EVIDENCE IN THE EUROPEAN UNION

Jesús López-Rodríguez
J. Andrés Faíña
Jose Lopez Rodríguez
De conformidad con la base quinta de la convocatoria del Programa de Estímulo a la Investigación, este trabajo ha sido sometido a evaluación externa anónima de especialistas cualificados a fin de contrastar su nivel técnico.

La serie DOCUMENTOS DE TRABAJO incluye avances y resultados de investigaciones dentro de los programas de la Fundación de las Cajas de Ahorros. 
Las opiniones son responsabilidad de los autores.
HUMAN CAPITAL ACCUMULATION AND GEOGRAPHY: EMPIRICAL EVIDENCE IN THE EUROPEAN UNION

Jesús López-Rodríguez, London School of Economics and University of A Coruña
J. Andrés Faíña, University of A Coruña and Jean Monnet Chair in European Industrial Economics
Jose Lopez Rodriguez, University of A Coruna

Abstract

This paper evaluates the role that geography plays in determining the spatial distribution of educational attainment levels among EU regions, based on an extension of the standard two sector (agriculture and manufacturing) Fujita et al. (1999) economic geography model. We provide evidence that, in the EU, educational attainment levels are higher in those regions with greater market access. This finding corroborates the theoretical predictions of the model and proves that remoteness is a penalty for the economic development and convergence of the European Union regions.

JEL classification: F12, F14, O10
Keywords: Economic Development, Economic Geography, Spatial structure

Authors’ address:

Jesus Lopez-Rodriguez            J. Andres Faíña            Jose Lopez Rodriguez
Research Scholar             Department of Economic Administration
Department of Economics, LSE Analysis and Business Administration
Houghton Street              University of A Coruña University of A Coruña
London WC2A 2AE, UK          Campus de Elviña s/n          Campus de Elviña s/n
j.lopez-rodriguez@lse.ac.uk   C.P.15.071               C.P.15.071
                              Coruña, Spain             A Coruña, Spain
                              fai@udc.es                jlopez@udc.es
                              Tel:+34981167050           Tel:+34981167050
                              Fax:+34981167070           Fax:+34981167070

Acknowledgements

This work has been partially done when the first author was a Visiting Scholar at the Department of Economics at LSE. I would like to thank Stephen Redding, Andrés Rodriguez-Pose, Antony Venables and David Jacho-Chávez for valuable comments. We are grateful for feedback received at the European Economic Association and European Regional Science Association conferences (Amsterdam 2005).
I would also like to thank comments made by a FUNCAS referee that contributed to improve the quality of the results. The usual disclaimer applies.
1. Introduction

In January 2003 the release of the 2nd intermediate report of the economic and social cohesion showed that in the year 2000 the regions with the highest GDP per head (PPS) accounting for 25% of the total population in the European Union were twice as rich as the regions with the lowest GDP per head (PPS) accounting for the 25% of the total population in the European Union. This ratio was the same in 1990. At the 10% percentile the situation was even worse, the ratio between the regions with the highest GDP per head levels and those with the lowest GDP per head levels was about 2.8 in 1990 and 2.6 in 2000. The persistence of such differences is surprising in light of the successive steps taken by the European Union towards higher levels of integration. The Single European Act, 1987 was conceived with the aim of fostering the integration process in the European Community. Integration was to be characterized by the twin dynamic of monetary union and regional expansion. The former process came about through the unification of European markets and the bases for monetary union which were set out in the EU Treaty, Maastricht (1992), and had the effect of intensifying integration and giving rise to economic and monetary union (EMU) by 1999. The deepening of economic integration was followed by a wide-scale reform in Regional Policy. Present Regional Policy has been shaped by the introduction of the Economic and Social Cohesion principle in the EEC Treaty by the Single European Act (SEA) in 1987. It was reinforced after the ratification of the EU Treaty and the creation of the new Cohesion Fund. There are a number of reasons which may prevented convergence of income levels such as sluggish technology diffusion, endowment disadvantages and trade costs. At this point New Economic Geography (NEG) has reached a theoretical consolidation as a theory that explains the emergence of a heterogeneous economic space. In NEG models agglomeration of economic activities and population bases on increasing returns to scale and transport costs. Firms that are far from large consumer markets pay greater transport costs and have less value added available to remunerate domestic factors of production. Although it is well documented the theoretical bases of NEG models, authors such as Neary (2001), Ottaviano (2002) and Head and Mayer (2003) pointed out that empirical research on NEG is lagging behind. Recently new empirical contributions came out (Brakman et al. 2000, 2002, Faina and Lopez-Rodriguez 2004, Hanson 1998, 2000, Niebuhr 2004, Mion 2003, Redding and Schott 2003, Redding and Venables 2004, Ross 2001).

The aim of the paper is to add to the empirical evidence on NEG. We investigate whether there exists a spatial educational attainment structure in the EU, i.e. whether there is a positive correlation between regional levels of educational attainment and distance from large consumer markets.

---

1 Data refer to EU15 countries
2 The figures of the 1st intermediate report on the economic and social cohesion comparing the years 1989 and 1999 for the 10% and 25% of population with the highest and lowest levels of GDP per head were the same as in the 2nd intermediate report on the economic and social cohesion.
3 With respect to the effectiveness of the European Union Regional Policy to boost regions whose development is lagging behind the opinions of the scholars are divergent (see Bastle et al. 2001, Boldrin and Canova 2001, Faina and Lopez-Rodriguez 2004, Rodriguez-Pose and Fratesi 2004)
6 See Krugman, 1991, 1992
Our study departs from an extension of the standard two sector (agriculture and manufacturing) Fujita et al. (1999) economic geography model carried out by Redding and Schott (2003) in which they allowed unskilled individuals to endogenously choose whether to invest in education. The authors demonstrate that there is a penalty of remoteness in terms of return to skill, thereby reducing incentives for investment in human capital accumulation. In Redding and Schott words “This penalty magnifies the effect that economic geography can have on cross-country per capita income; increasing a country’s relative trade costs not only reduces contemporaneous factor rewards, but also lowers gross domestic product by suppressing human capital accumulation and decreasing the supply of high-income skilled workers”

Our investigation complements the work carried out by Redding and Schott (2003). It differs from them in geographical focus (European Union regions rather than World countries) and in the fact that we use and ad hoc measure of market access instead of using the theory-based measure.

The rest of the paper is structured as follows: Section 2 contains a brief description of the theoretical model that constitutes the theoretical framework of the empirical analysis. In section 3 contains the empirical framework, data and regional system use in our estimations. The results of the regression analysis are presented in section 4. Section 5 contains the main concluding remarks. In section 6 are the main references used.

2. The Theoretical Model

In NEG models, the interaction of transport costs and increasing returns to scale generates demand linkages and serves as explanation for agglomeration. Agglomeration is caused by a circular relationship in which the spatial concentration of manufacturing both creates and follows market access. In krugman’s (1991a) words, circular causation a la Myrdal is present because these two effects reinforce each other: “manufactures production will tend to concentrate where there is a large market, but the market will be large where manufactures of production is concentrated”. These forces that are at work in any multiregional economy can be studied within a relatively simple general equilibrium model of monopolistic competition developed by Krugman (1991b), which has come to be known as the core-periphery model7. Krugman’s theoretical research on NEG has triggered a plethora of contributions8, which have been surveyed by Ottaviano and Puga (1998). Most recently a synthesis of the existing theoretical research on NEG can be found in Fujita et al. (1999) and Fujita and Thisse (2002).

The present analysis is based on an extension of the standard two sector (agriculture and manufacturing) Fujita et al. (1999) economic geography model carried out by Redding and Schott (2003) by introducing human capital accumulation. The structure of Redding and Schott’s model is as follows. The economic space is made of \(i \in \{1, \ldots, R\}\) countries. Each country is endowed with \(iL\) consumers. Consumers have one unit of labour which is supplied inelastically with zero disutility and consumers choose endogenously whether or not to invest in becoming skilled. Preferences are identical across all consumers and described by a Cobb-Douglas utility function,

\[
U_j = A_j^{1-\mu} M_j^{\mu}, \quad 0 < \mu < 1
\]

7 An earlier analysis that anticipated several aspects of Krugman’s work was developed by Faini (1984). Ideas close to economic geography have already appeared in Krugman (1979) but were not fully worked out.

Where \( \mu \) is the share of expenditure on manufactures, \( A \) is the quantity of agricultural good consumed and \( M \) is a composite of symmetric product varieties given by:

\[
M_j = \left[ \sum_{i=1}^{n} m^C_{i,j}(x) \right]^{\frac{\sigma-1}{\sigma}} = \left[ \sum_{i=1}^{n} \left( n_i \cdot m^C_{i,j} \right) \right]^{\frac{\sigma-1}{\sigma}}
\]

(2)

We use \( j \) to denote a country that is demanding or importing a good and \( i \) to denote a country that is producing or exporting a good. \( \sigma \) is the constant elasticity of substitution between any pair of varieties (consumers have love for variety, with increasing \( \sigma \), the substitutability among varieties rises, thus the desire to spread consumption over manufactured goods declines), \( n_i \) denotes the number of varieties produced in country \( i \) and \( m^C_{i,j} \) denotes the amount of each variety produced in country \( i \) for final consumption in country \( j \). The second equation exploits the fact that, in equilibrium all the products produced in country \( i \) are demanded by country \( j \) in the same quantity, so we can remove the index \( z \) and rewrite the integral as a product.

Dual to manufacturing goods consumption index (\( M_j \)) is a price index (\( G_j \)) defined over the prices of individual varieties produced in \( i \) and sold in \( j \) (i.e.) \( P_{i,j} \).

\[
G_j = \left[ \sum_{i=1}^{n} \left( P_{i,j} \right) \right]^{\frac{1}{\sigma}} = \left[ \sum_{i=1}^{n} \left( n \cdot P_{i,j} \right) \right]^{\frac{1}{\sigma}} = \left[ S_A \right]^{\frac{1}{\sigma}}
\]

(3)

Equation (3) can also be interpreted as a measure of the country’s overall access to sources of supply-supplier access (\( S_A \)).

The agricultural sector produces a homogeneous good under constant returns to scale

\[
Y_i = \theta_i (S_i^\phi) (L_i^{1-\phi}), \quad 0 < \phi < 1
\]

(4)

\( Y_i \) denotes the output in the agricultural sector which is endowed with \( L_i \) unskilled workers and with \( S_i \) skilled workers. \( \theta_i \) stands for agricultural productivity. Here, we assume that the output of the agricultural sector is costlessly traded between any two countries.

The manufacturing sector produces a differentiated good according to an increasing returns to scale technology such as the production of each variety requires primary factors of production (skilled and unskilled labour) and goods from the manufacturing sector.

\footnote{Redding and Schott (2003) also relax this assumption considering the possibility of trade costs involved in shipped the agricultural good.}
sector (intermediate inputs). The costs that a firm in country $i$ incurs to produce the differentiated good can be given by:

$$\Gamma_i = (w_i^s)^\alpha (w_i^u)^\beta G_i^{(1-\alpha-\beta)} c_i [F + x_i],$$  \hspace{1cm} (5)

where $w_i^s$ is the wage of skilled workers (input share $\alpha$), $w_i^u$ is the wage of unskilled workers (input share $\beta$), $G_i$ is the price index for manufactures (input share $1-\alpha-\beta$), $c_i$ is a constant marginal input requirement, $F$ is a fixed input requirement and $x_i = \sum_{j=1}^{K} x_{ij}$ is the total output of the firm produced for all markets.

Manufactured goods are trade among countries incurring iceberg costs, i.e. a fraction of any good shipped from location $i$ to location $j$ melts away so in order to arrive at location $j$ with one unit of good $T_{ij}^M > 1$ units must be shipped.

In order to introduce endogenous human capital accumulation in the model, Redding and Schott (2003) assume that one unit of unskilled labour in country $i$ ($iz_i$) can be converted in one unit of skilled labour by incurring in a fixed cost of education $h_i$ units of unskilled labour where $h_i$ is an inverse measure of the extend of public provision education and $a(z)$ is the individual ability that has an upper and lower bounds determined by human biology $[\underline{a}, \overline{a}]$. Associated with this distribution of ability we can defined a probability density function $\lambda(a)$ and a cumulative distribution of ability $\Lambda(a) = \int_{\underline{a}}^{a} \lambda(a) da$.

Taking into account the features above mentioned the decision to become educated will be given by:

$$w_i^s - w_i^u \geq \frac{h_i}{a(z)} w_i^u$$  \hspace{1cm} (6)

This expression tells us that there is incentives to invest in education when the differential wage between skilled workers and unskilled ones exceed the educational costs. Moreover from equation (6) it can be obtained the equilibrium supply of skills, i.e., the critical value for $a$ ($a^*_i$) such that if $a(z) \geq a^*_i(z^*)$ all individuals choose to become skilled:

$$a^*_i = \frac{h_i}{\frac{w_i^s}{w_i^u} - 1}$$  \hspace{1cm} (7)

For a rigorous analysis see Redding and Schott (2003)
The worker with ability \(a^*_i\) is indifferent between becoming skilled and remaining unskilled, so this equation is called *skill indifference condition* \((S)\). The equilibrium conditions on the consumer’s side are obtained by maximizing their utility subject to their budget constraint. If we denote by \(E_j^C\) total consumer expenditure on manufacturing goods in country \(j\), its demand for each product is (Applying Shephard’s lemma on the price index\(^{11}\))

\[
m_j^C = (P_j^M)^{-\sigma} E_j^C G_j^{\sigma - 1} \tag{8}
\]

The term \(E_j^C G_j^{\sigma - 1}\) is a measure of demand in the importing country \(j\) termed *market capacity* \((m_j^C = E_j^C G_j^{\sigma - 1})\) and comprised of total expenditure on manufacturing goods in market \(j\) \((E_j^C)\) as well as the number of competing firms and the prices they charge as summarized in the manufacturing price index \((G_j)\).

With respect to the producers’ equilibrium, in the agricultural sector profit maximization imply that price equals unit costs of production:

\[
P_i^Y = 1 = \frac{1}{\varphi_i^F} (w_i^S)^{\phi} (w_i^U)^{1-\phi} \tag{9}
\]

where the output of the agricultural good is chosen as the numeraire, and thus \(P_i^Y = 1\) for all \(i\). In the manufacturing sector, the profit function of a representative country \(i\) firm is:

\[
\Pi_i = \sum_{j=1}^{\mathcal{M}} \frac{P_{ij}^M x_{ij}}{T_j^M} - (w_i^S)^{\alpha} (w_i^U)^{\beta} G_j^{1-\alpha-\beta} c_i (F + x_i) \tag{10}
\]

Because varieties are equally weighed in the utility function, the equilibrium price is the same across all firms located in region \(i\). Solving the first order condition yields the common equilibrium price:

\[
P_i^M = \left(\frac{\sigma}{\sigma - 1}\right) (w_i^S)^{\alpha} (w_i^U)^{\beta} G_i^{1-\alpha-\beta} c_i \tag{11}
\]

\(^{11}\) See also Dixit and Stiglitz (1977)
This means that the firm uses a relative mark-up over marginal costs. Substituting (11) into the profit function leads to:

$$\Pi_i = \left( \frac{p_i^M}{\sigma} \right) [x_i - (\sigma - 1)F]$$  \hspace{1cm} (12)

Under free entry, profits are zero, and thus equilibrium output of a firm is a constant given by:

$$x = \frac{\sigma - 1}{F}$$  \hspace{1cm} (13)

The price needed to sell this many units satisfies\(^{12}\) (using demand function (4))

$$(p_i^M) = \frac{1}{x} \sum_{j=1}^{R} E_j G_j^{\sigma-1} (T_j^M)^{1-\sigma}$$  \hspace{1cm} (14)

Introducing the equilibrium prices (11) in equation (14) we obtain the following zero-profit condition:

$$\left[ \left( \frac{\sigma}{\sigma - 1} \right) (w_j^y)^{\sigma} (w_j^U)^{\beta} G_j^{1-\sigma-\beta} c_j \right]^{\sigma} = \frac{1}{x} \sum_{j=1}^{R} E_j G_j^{\sigma-1} (T_j^M)^{1-\sigma}$$  \hspace{1cm} (15)

This is the so-called nominal wage equation which is point of departure of most studies that investigate the existence of a spatial wage structure for different cross sections. According to equation (15), the nominal wage level in region \(i\) depends on a weighted sum of purchasing power in all accessible regions \(j\), whereby the weighting scheme is a function declining with increasing distance between locations \(i\) and \(j\). This sum we will refer to as the “market access” of country \(i\) (\(MA_i\)). As Hanson (2000) notes, equation (15) can be thought of as a spatial labour demand function in an economy with perfect labour mobility. Labour demand and wages increase with income of neighbouring regions and decline with rising transport costs to these locations. The nominal wage equation represents one of the main propositions emerging from NEG models mentioned by Head and Mayer (2003): access advantages raise local factor prices. More precisely, production sites with good access to major markets because of relatively low trade costs tend to reward their production factors with higher wage and land rentals. Moreover, the equation resembles the market potential concept introduced

\(^{12}\) The transport cost term \((T_{i,j})\) enters with the exponent \((1 - \sigma)\) and not \(\sigma\) because total shipments to market \(j\) are \(T_{i,j}\) times quantities consumed
by Harris (1954). The market potential concept states that the attractiveness of a region as a production site depends on its access to markets.

The nominal wage equation can be rewritten as:

\[(w_i^s)^\alpha (w_i^u)^\beta = \xi \frac{1}{c_i} (MA_i)^{\frac{1}{\sigma}} G_i^{(\alpha + \beta - 1)} \quad (16)\]

by making use of equation (3) we can express the skilled and unskilled wages as a function of the market access and supplier access.

\[(w_i^s)^\alpha (w_i^u)^\beta = \xi \frac{1}{c_i} (MA_i)^{\frac{1}{\sigma}} (SA_i)^{\frac{1-\alpha-\beta}{\sigma-1}} \quad (17)\]

where the constant \(\xi\) on the right-hand side combines constants from the equation (15). Equations (9) and (17) combined together give us the equilibrium wages for skilled and unskilled workers. Taking logs and differentiating equations (9) and (17) and combining them with the skill indifference condition-equation (7)-, it is obtained the equilibrium relationship between geographical location and endogenous human capital investments.

\[0 = \phi \frac{dw_i^s}{w_i^s} + (1 - \phi) \frac{dw_i^u}{w_i^u} \quad (18)\]

\[\alpha \frac{dw_i^s}{w_i^s} + \beta \frac{dw_i^u}{w_i^u} = \frac{1}{\sigma} \frac{dMA_i}{MA_i} + \frac{(1 - \alpha - \beta)}{(\sigma - 1)} \frac{dSA_i}{SA_i} \quad (19)\]

Taking into account equations (18) and (19) it can be shown that if the equilibrium market access \((MA_i)\) and supplier access \((SA_i)\) fall and if the manufacturing sector is skill-intensive relative to the agricultural sector and the country remains incompletely specialized (this will occur for relatively high values of trade costs), the new equilibrium must be characterized by a lower relative wage of skilled workers and therefore by using the skill indifference condition this new equilibrium implies a higher critical level of ability \(a_i^*\) above which individuals become skilled and a reduced (increased) supply of skilled (unskilled) workers.

The intuitive explanation is based on that the fall in the market and supplier access modifies the initial equilibrium conditions in the manufacturing sector which experiences a fall in size. This reduction in size releases more skilled labour than is demanded initially in the agricultural sector. To go back to the equilibrium point, the nominal skilled wage has to be lower and the nominal unskilled wage higher and therefore the relative wage of skilled workers is lower. Taking into account equation (7), a lower relative wage of skilled workers reduces the incentive to invest in education. Therefore, there is a positive (negative) correlation between market access and number of skilled (unskilled) workers.
3. Empirical framework, data and regional system

The results that are obtained from the theoretical model can be tested by using the following regression equation:

\[
\ln(EA_{it}) = \alpha_{0t} + \alpha_{it} \ln(MA_{it}) + \varepsilon_{it}
\]  

Equation (20) allows us to check if there is a spatial educational attainment structure in the EU, i.e. whether there is a positive correlation between medium and high levels of educational attainment and distance from large consumer markets, i.e. if high market access locations have relatively high levels of education. 

\(EA_{it}\) stands for levels of educational attainment for different countries/regions and periods of time. In our analysis, the dependent variable in the regression analysis of European regions is the log of educational attainment. The corresponding data were taken from the Eurostat Regio databank -first and second intermediate reports on economic and social cohesion (dates of release 30-01-2002 and 30-01-2003) and the third report on economic and social cohesion (date of release 2004)-. The educational attainment is a three-level measure defined as persons aged 25-59 (as a % of total) with low, medium or high levels of education. The dependent variable is given for 203 European Union NUTS2 regions. The right-hand side of the equation contains the market access variable, a constant and the random disturbance. With respect to the market access measure \(MA_{it}\), we built an ad hoc measure of it, that is the population potential computed for every single place in the European space. The concept of population potential must be understood as the force or attraction which the population centre \(j\) would exert on an inhabitant located at location \(i\) in geographical space and conditioned according to the distance between them, \(T_{ij}\). Therefore, this measure shows the influence each place exerts on all other places and in this sense it measures the proximity of a place to other places. This measure of market access is quite similar to the market potential concept proposed by Harris (1954) and also there is a natural link with the concept of demand cones due to Lösch (1954). Population potentials at a given location represent an index of the aggregate market potential from the whole structure of population weighing the number of inhabitants by their distance to this location. In order to make the computations we have to say that it is not possible to consider all the points within a given territory, which is why the practical computation of the population potentials is carried out by using a dot or grid “net”. This net which, is placed over specific \(j\) places, defines a finite and manageable set of nodes for the calculations. The potential indices are calculated by going through each node on the net and assigning to it a corresponding “potential” value, that is, the value of its own population plus the inhabitants of each and every other node divided by the distance separating them to the original node.

The calculations were carried out in the following way: for each “\(i\)” node in the net we add the population of each centre divided by:

The distance \(T_{ij}\), measured in kilometers, if it is more than 1, or

One, if the distance is less than 1.
To this end, an algorithm or “loop” which goes through the whole of the net \( (i) \) is designed to complete the whole of the space and is computed in order to be able to compile the indices. The population data we used have been obtained from the statistics information service of the European Commission, Eurostat, and the cartographic data from GISCO. From the group of urban centers with more than twenty thousand inhabitants the potential index of each point of the net was calculated in ARC/INFO and then they have been plotted in a map by using ARC MAP. The next step in our computations was to assign a market access value to each of the NUTS2 regions of the European Union in order to have a comparable relationship between levels of educational attainment and these population potentials, i.e. the same geographical coverage. To this end the market access value assigned to each of the NUTS2 regions in the European Union is based on a weighted aggregation of the points’ population potential that belong to a particular region. Before analyzing the regression results it is interesting to graphically see a map that displays a classification in five levels of our market access measure within the EU15.\(^{13}\)

The value of the market access measure (population potential) is reflected in the relative shade of the colour used, that is, the darker the shade, the higher the market access (population potential) and visa versa. The map reflects a concentric distribution of the market access measure, which has its centre in an area in which the values are the highest, an area that is commonly known as the Golden triangle (Greater Manchester-London-Paris and the Rhur Valley). This area is surrounded by successive envelopes of decreasing market access values, which eventually reach the Atlantic periphery where the values are lowest.

\(^{13}\) For a detailed analysis of the spatial distribution of population potentials in the European space see (Faiña et al., 2001)
4. Economic Geography and Educational Attainment Levels: Empirical results

In this section we test econometrically expression (20), i.e. whether there is a positive correlation between human capital investments and market access. Consistent with the model, we provide evidence that educational attainment is higher in those regions that have greater market access. Figures 1, 2 and 3 plot medium, high and medium+high educational attainments against market access for year 2000. It is clear from these figures that the relationship between regional levels of medium, high and medium+high educational attainments and regional market access are in line with the predictions of the model. The relationship is robust and is not due to the influence of a few individual regions.
Figure 2: Regional market access vs human capital
EU15, Year 2000
(High Educational Attainments)

Figure 3: Regional market access vs human capital
EU15, Year 2000
(Medium+High Educational Attainments)
Tables 1-3 present the results of our econometric estimations for 203 EU NUTS2 regions. We regress log medium educational attainment, log high educational attainment and log medium-high educational attainment on log ad hoc measure of market access using OLS. The coefficients on market access are significant and the signs correspond with theoretical expectations.

Table 1: OLS Regression of Medium Educational Attainment on Regional Market Access, Year 2000
Dependent Variable: Log Medium Educacional Attainment
Included observations: 203 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-6.809588</td>
<td>1.887801</td>
<td>-3.607154</td>
<td>0.0004</td>
</tr>
<tr>
<td>Log (MA)</td>
<td>0.998549</td>
<td>0.143535</td>
<td>6.956853</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared 0.194059 Mean dependent var 6.316829
Adjusted R-squared 0.190049 S.D. dependent var 0.956941
S.E. of regression 0.861221 Akaike info criterion 2.548873
Sum squared resid 149.0821 Schwarz criterion 2.581515
Log likelihood -256.7106 F-statistic 48.39781
Prob(F-statistic) 0.000000

Table 2: OLS Regression of High Educational Attainment on Regional Market Access, Year 2000
Dependent Variable: Log High Educational Attainment
Included observations: 203 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-6.262878</td>
<td>1.921406</td>
<td>-3.259529</td>
<td>0.0013</td>
</tr>
<tr>
<td>LNV</td>
<td>0.903618</td>
<td>0.146090</td>
<td>6.185365</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared 0.159905 Mean dependent var 5.615624
Adjusted R-squared 0.155726 S.D. dependent var 0.953973
S.E. of regression 0.876552 Akaike info criterion 2.584162
Sum squared resid 154.4370 Schwarz criterion 2.616804
Log likelihood -260.2924 F-statistic 38.25875
Prob(F-statistic) 0.000000

Table 3 OLS Regression of Medium+High Educational Attainment on Regional Market Access, Year 2000
Dependent Variable: Log Medim+High Educational Attainment
Included observations: 203 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-5.913805</td>
<td>1.806912</td>
<td>-3.272879</td>
<td>0.0013</td>
</tr>
<tr>
<td>LN (MA)</td>
<td>0.963220</td>
<td>0.137384</td>
<td>7.011127</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

R-squared 0.196501 Mean dependent var 6.478192
Adjusted R-squared 0.192504 S.D. dependent var 0.917330
S.E. of regression 0.824320 Akaike info criterion 2.461287
Sum squared resid 136.5801 Schwarz criterion 2.493929
Log likelihood -247.8206 F-statistic 49.15591
Prob(F-statistic) 0.000000

These results show that almost 20% of the variation in regional levels of secondary and tertiary education is explained by market access. However, the models given in tables 1-3 are marked by outlying observations. The outlying regions do not correspond with the spatial educational attainment structure determined by the majority of the observations.
Outliers will seriously affect the coefficients estimates, if they are influential leverage points, i.e. outlying observations with regard to our measure of market access. In order to control for effects of outlying observations, dummy variables for the outliers are introduced. The most significant outliers are the regions of Bruxelles and Inner London. Moreover, the regions of Wien, Andalucia and Denmark are outliers as well. To control for the latter observations, regional dummies were included. Including dummies for outlying regions does alter the estimates considerably (see tables 4-6). In particular, the coefficient of the market access increases in all the regressions. According to the results, market access has a positive effect on the educational attainment levels of the European Union regions. This effect is reinforced when dummy variables are included in the regressions. The models presented in tables 4 and 6 explain around 30% of the spatial variation in the educational attainment levels in the European Union.14

Table 4: OLS Regression of Medium Educational Attainment on Regional Market Access and Dummy Variables, Year 2000
Dependent Variable: Log Medium Educacional Attainment
Included observations: 203 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-12.86948</td>
<td>2.162430</td>
<td>-5.951395</td>
<td>0.0000</td>
</tr>
<tr>
<td>LN(MA)</td>
<td>1.461293</td>
<td>0.164746</td>
<td>8.869987</td>
<td>0.0000</td>
</tr>
<tr>
<td>DAT13</td>
<td>-1.549939</td>
<td>0.847773</td>
<td>-1.828249</td>
<td>0.0690</td>
</tr>
<tr>
<td>DBE1</td>
<td>-2.765784</td>
<td>0.840810</td>
<td>-3.289428</td>
<td>0.0012</td>
</tr>
<tr>
<td>DUKI1</td>
<td>-3.568617</td>
<td>0.933999</td>
<td>-3.820793</td>
<td>0.0002</td>
</tr>
<tr>
<td>DDK</td>
<td>1.916596</td>
<td>0.813188</td>
<td>2.356891</td>
<td>0.0194</td>
</tr>
<tr>
<td>DES61</td>
<td>1.275771</td>
<td>0.816372</td>
<td>1.562733</td>
<td>0.1197</td>
</tr>
</tbody>
</table>

R-squared 0.303641  Mean dependent var 6.316829
Adjusted R-squared 0.282324  S.D. dependent var 0.956941
S.E. of regression 0.810680  Akaike info criterion 2.451988
Sum squared resid 128.8117  Schwarz criterion 2.566237
Log likelihood -241.8768  F-statistic 14.24399
Prob(F-statistic) 0.000000

14The market access variable remains statistically significant at conventional critical values including control variables thought to be important in the determination of educational levels (Gross Value Added Per capita, employment in high tech sectors and backwardness). Proofs can be sent by the authors upon request.
Table 5: OLS Regression of High Educational Attainment on Regional Market Access and Dummy Variables, Year 2000
Dependent Variable: Log High Educational Attainment
Included observations: 203 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-9.788321</td>
<td>2.273357</td>
<td>-4.305667</td>
<td>0.0000</td>
</tr>
<tr>
<td>LN (MA)</td>
<td>1.171900</td>
<td>0.173197</td>
<td>6.766288</td>
<td>0.0000</td>
</tr>
<tr>
<td>DAT13</td>
<td>-1.517650</td>
<td>0.891261</td>
<td>-1.702812</td>
<td>0.0902</td>
</tr>
<tr>
<td>DBE1</td>
<td>-1.162368</td>
<td>0.883942</td>
<td>-1.314983</td>
<td>0.1901</td>
</tr>
<tr>
<td>DUK1</td>
<td>-1.639608</td>
<td>0.981911</td>
<td>-1.669814</td>
<td>0.0966</td>
</tr>
<tr>
<td>DDK</td>
<td>1.860134</td>
<td>0.854903</td>
<td>2.175843</td>
<td>0.0308</td>
</tr>
<tr>
<td>DES61</td>
<td>2.204463</td>
<td>0.858250</td>
<td>2.568557</td>
<td>0.0110</td>
</tr>
</tbody>
</table>

R-squared 0.225568  Mean dependent var 5.615624
Adjusted R-squared 0.201861  S.D. dependent var 0.953973
S.E. of regression 0.852266  Akaike info criterion 2.552038
Sum squared resid 142.3661  Schwarz criterion 2.666287
Log likelihood -252.0319  F-statistic 9.514774
Prob(F-statistic) 0.000000

Table 6: OLS Regression of Medium+High Educational Attainment on Regional Market Access and Dummy Variables, Year 2000
Dependent Variable: Log Medium+High Educational Attainment
Included observations: 203 after adjusting endpoints

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-10.86047</td>
<td>2.099596</td>
<td>-5.172646</td>
<td>0.0000</td>
</tr>
<tr>
<td>LN (MA)</td>
<td>1.340556</td>
<td>0.159959</td>
<td>8.380634</td>
<td>0.0000</td>
</tr>
<tr>
<td>DAT13</td>
<td>-1.512465</td>
<td>0.823138</td>
<td>-1.837437</td>
<td>0.0677</td>
</tr>
<tr>
<td>DBE1</td>
<td>-2.053054</td>
<td>0.816378</td>
<td>-2.514832</td>
<td>0.0127</td>
</tr>
<tr>
<td>DUK1</td>
<td>-2.740392</td>
<td>0.906859</td>
<td>-3.021849</td>
<td>0.0028</td>
</tr>
<tr>
<td>DDK</td>
<td>1.866885</td>
<td>0.789559</td>
<td>2.364466</td>
<td>0.0190</td>
</tr>
<tr>
<td>DES61</td>
<td>1.676735</td>
<td>0.792650</td>
<td>2.115353</td>
<td>0.0357</td>
</tr>
</tbody>
</table>

R-squared 0.285602  Mean dependent var 6.748192
Adjusted R-squared 0.263733  S.D. dependent var 0.917330
S.E. of regression 0.787124  Akaike info criterion 2.393012
Sum squared resid 121.4346  Schwarz criterion 2.507261
Log likelihood -235.8908  F-statistic 13.05947
We also report the estimates of an ordered probit model (table 7). The dependent variable (Educational Attainment Level) was ranked for each region given to it the values 1 (low educational attainment), 2 (medium educational attainment) or 3 (high educational attainment) taking into account the relative importance of the educational attainment shares in each region.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LN (MA)</td>
<td>1.189026</td>
<td>0.228852</td>
<td>5.195600</td>
</tr>
</tbody>
</table>

Limit Points

<table>
<thead>
<tr>
<th>Limit Points</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIMIT_2:C(2)</td>
<td>15.17475</td>
<td>2.994245</td>
<td>5.067973</td>
<td>0.0000</td>
</tr>
<tr>
<td>LIMIT_3:C(3)</td>
<td>18.14431</td>
<td>3.127138</td>
<td>5.802209</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Ordered Probit Models, the sign of the coefficient shows the direction of the change in the probability of falling in the endpoint rankings, in our case (Educational Attainment Level=1 or Educational Attainment Level=3) when Market Access changes. Probability of Educational Attainment level=1 changes in the opposite direction of the sign of the estimated coefficient and Probability of Educational Attainment level=3 changes in the same direction. So the results show that the probability that the educational attainment level of a region is high increases with market access. The coefficient estimated is statistically significant at the conventional critical values.

Our results are in line with those obtained by Redding and Schott (2003) for a world sample of countries. In their estimations market access itself explained 23% of the variation in educational attainment (105 countries) and excluding from the sample OECD countries, US, Japan and Belgium (66 countries) the explanatory power of the regression raised to 26%.

These results shed new light to the pioneering work initiated by Redding and Schott (2003), showing that at the EU level there is a positive correlation between countries’ human capital investments and market access. As Redding and Schott (2003) pointed out fruitful avenues for future research include the analysis of the relationship between changes in educational attainment and changes in market access within countries and the exploitation of exogenous changes in market access associated with changes in policy regimes.
5. Conclusions

In this paper, we analyse the relationship between an *ad hoc* measure of market access (Population potential) and the levels of educational attainment in the European Union regions for the year 2000. Consistent with the predictions of the theoretical model, we provide empirical evidence of a spatial educational attainment structure in the EU, i.e. a positive correlation between regional medium and high levels of educational attainment and distance from large consumer markets. The inclusion of dummy variables alter the coefficient of market access considerably, changing from values around but below to 1 to values over 1.17. The augmented models explain around 30% of the spatial variation in the educational attainment levels in the European Union regions.

Taking into account that human capital accumulation is a key factor for regional development and to promote convergence among EU regions and the results of this analysis suggesting that there is a penalty of remoteness for human capital accumulation, one obvious policy implication is that the outlying regions in the EU should make bigger efforts to improve the quality of their infrastructures trying to reduce distance to the main centres of economic activity. An important role in this sense has been played by the European Union Regional Policy since its institutionalization (1989), devoting an important part of its resources to objective 1 regions (most of them in the outskirts of the EU and so facing the penalty of the remoteness) throughout its three programming periods (Delors I and II packages and Agenda 2000). The majority of resources where channelled to improvements in infrastructure, human capital and aids to production sectors.
6. References


<table>
<thead>
<tr>
<th>Número</th>
<th>Título</th>
<th>Autor(es)</th>
</tr>
</thead>
<tbody>
<tr>
<td>159/2000</td>
<td>Participación privada en la construcción y explotación de carreteras de peaje</td>
<td>Ginés de Rus, Manuel Romero y Lourdes Trujillo</td>
</tr>
<tr>
<td>160/2000</td>
<td>Errores y posibles soluciones en la aplicación del <em>Value at Risk</em></td>
<td>Mariano González Sánchez</td>
</tr>
<tr>
<td>161/2000</td>
<td>Tax neutrality on saving assets. The spahish case before and after the tax reform</td>
<td>Cristina Ruza y de Paz-Curbera</td>
</tr>
<tr>
<td>163/2000</td>
<td>El control interno del riesgo. Una propuesta de sistema de límites riesgo neutral</td>
<td>Mariano González Sánchez</td>
</tr>
<tr>
<td>164/2001</td>
<td>La evolución de las políticas de gasto de las Administraciones Públicas en los años 90</td>
<td>Alfonso Utrilla de la Hoz y Carmen Pérez Esparrells</td>
</tr>
<tr>
<td>165/2001</td>
<td>Bank cost efficiency and output specification</td>
<td>Emili Tortosa-Ausina</td>
</tr>
<tr>
<td>166/2001</td>
<td>Recent trends in Spanish income distribution: A robust picture of falling income inequality</td>
<td>Josep Oliver-Alonso, Xavier Ramos y José Luis Raymond-Bara</td>
</tr>
<tr>
<td>167/2001</td>
<td>Efectos redistributivos y sobre el bienestar social del tratamiento de las cargas familiares en</td>
<td>Nuria Badenes Plá, Julio López Laborda, Jorge Onrubia Fernández</td>
</tr>
<tr>
<td></td>
<td>el nuevo IRPF</td>
<td></td>
</tr>
<tr>
<td>168/2001</td>
<td>The Effects of Bank Debt on Financial Structure of Small and Medium Firms in some European</td>
<td>Mónica Melle-Hernández</td>
</tr>
<tr>
<td></td>
<td>Countries</td>
<td></td>
</tr>
<tr>
<td>169/2001</td>
<td>La política de cohesión de la UE ampliada: la perspectiva de España</td>
<td>Ismael Sanz Labrador</td>
</tr>
<tr>
<td>170/2002</td>
<td>Riesgo de liquidez de Mercado</td>
<td>Mariano González Sánchez</td>
</tr>
<tr>
<td>171/2002</td>
<td>Los costes de administración para el afiliado en los sistemas de pensiones basados en cuentas</td>
<td>José Enrique Devesa Carpio, Rosa Rodríguez Barrera, Carlos Vidal Meliá</td>
</tr>
<tr>
<td></td>
<td>de capitalización individual: medida y comparación internacional.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>y propuestas de metodología para la explotación de la información de los ingresos y el gasto.</td>
<td></td>
</tr>
<tr>
<td>173/2002</td>
<td>Modelos paramétricos y no paramétricos en problemas de concesión de tarjetas de credito.</td>
<td>Rosa Puertas, María Bonilla, Ignacio Olmeda</td>
</tr>
</tbody>
</table>
174/2002 Mercado único, comercio intra-industrial y costes de ajuste en las manufacturas españolas. José Vicente Blanes Cristóbal

175/2003 La Administración tributaria en España. Un análisis de la gestión a través de los ingresos y de los gastos. Juan de Dios Jiménez Aguilera, Pedro Enrique Barrilao González


177/2003 Effects of ATMs and Electronic Payments on Banking Costs: The Spanish Case. Santiago Carbó Valverde, Rafael López del Paso, David B. Humphrey

178/2003 Factors explaining the interest margin in the banking sectors of the European Union. Joaquín Maudos y Juan Fernández Guevara

179/2003 Los planes de stock options para directivos y consejeros y su valoración por el mercado de valores en España. Mónica Melle Hernández


181/2003 The Euro effect on the integration of the European stock markets. Mónica Melle Hernández

182/2004 In search of complementarity in the innovation strategy: international R&D and external knowledge acquisition. Bruno Cassiman, Reinhilde Veugelers

183/2004 Fijación de precios en el sector público: una aplicación para el servicio municipal de suministro de agua. Mª Ángeles García Valiñas

184/2004 Estimación de la economía sumergida es España: un modelo estructural de variables latentes. Ángel Alañón Pardo, Miguel Gómez de Antonio

185/2004 Causas políticas y consecuencias sociales de la corrupción. Joan Oriol Prats Cabrera

186/2004 Loan bankers’ decisions and sensitivity to the audit report using the belief revision model. Andrés Guiral Contreras and José A. Gonzalo Angulo

187/2004 El modelo de Black, Derman y Toy en la práctica. Aplicación al mercado español. Marta Tolentino García-Abadillo y Antonio Díaz Pérez

188/2004 Does market competition make banks perform well?. Mónica Melle

189/2004 Efficiency differences among banks: external, technical, internal, and managerial. Santiago Carbó Valverde, David B. Humphrey y Rafael López del Paso
190/2004 Una aproximación al análisis de los costes de la esquizofrenia en españa: los modelos jerárquicos bayesianos
F. J. Vázquez-Polo, M. A. Negrín, J. M. Cavasés, E. Sánchez y grupo RIRAG

191/2004 Environmental proactivity and business performance: an empirical analysis
Javier González-Benito y Óscar González-Benito

192/2004 Economic risk to beneficiaries in national defined contribution accounts (NDCs)
Carlos Vidal-Meliá, Inmaculada Domínguez-Fabian y José Enrique Devesa-Carpio

193/2004 Sources of efficiency gains in port reform: non parametric malmquist decomposition tfp index for Mexico
Antonio Estache, Beatriz Tovar de la Fé y Lourdes Trujillo

194/2004 Persistencia de resultados en los fondos de inversión españoles
Alfredo Ciriaco Fernández y Rafael Santamaría Aquilué

195/2005 El modelo de revisión de creencias como aproximación psicológica a la formación del juicio del auditor sobre la gestión continuada
Andrés Guiral Contreras y Francisco Esteso Sánchez

196/2005 La nueva financiación sanitaria en España: descentralización y prospectiva
David Cantarero Prieto

197/2005 A cointegration analysis of the Long-Run supply response of Spanish agriculture to the common agricultural policy
José A. Mendez, Ricardo Mora y Carlos San Juan

198/2005 ¿Refleja la estructura temporal de los tipos de interés del mercado español preferencia por la liquidez?
Magdalena Massot Perelló y Juan M. Nave

199/2005 Análisis de impacto de los Fondos Estructurales Europeos recibidos por una economía regional: Un enfoque a través de Matrices de Contabilidad Social
M. Carmen Lima y M. Alejandro Cardenete

200/2005 Does the development of non-cash payments affect monetary policy transmission?
Santiago Carbó Valverde y Rafael López del paso

201/2005 Firm and time varying technical and allocative efficiency: an application for port cargo handling firms
Ana Rodríguez-Álvarez, Beatriz Tovar de la Fé y Lourdes Trujillo

202/2005 Contractual complexity in strategic alliances
Jeffrey J. Reuer y Africa Ariño

203/2005 Factores determinantes de la evolución del empleo en las empresas adquiridas por opa
Nuria Alcalde Fradejas y Inés Pérez-Soba Aguilar

Elena Olmedo, Juan M. Valderas, Ricardo Gimeno and Lorenzo Escot
205/2005 Precio de la tierra con presión urbana: un modelo para España
Esther Decimavilla, Carlos San Juan y Stefan Sperlich

206/2005 Interregional migration in Spain: a semiparametric analysis
Adolfo Maza y José Villaverde

207/2005 Productivity growth in European banking
Carmen Murillo-Melchor, José Manuel Pastor y Emili Tortosa-Ausina

Santiago Carbó Valverde, David B. Humphrey y Rafael López del Paso

209/2005 La elasticidad de sustitución intertemporal con preferencias no separables intratemporalmente: los casos de Alemania, España y Francia.
Elena Márquez de la Cruz, Ana R. Martínez Cañete y Inés Pérez-Soba Aguilar

210/2005 Contribución de los efectos tamaño, book-to-market y momentum a la valoración de activos: el caso español.
Begoña Font-Belaire y Alfredo Juan Grau-Grau

211/2005 Permanent income, convergence and inequality among countries
José M. Pastor and Lorenzo Serrano

212/2005 The Latin Model of Welfare: Do 'Insertion Contracts' Reduce Long-Term Dependence?
Luis Ayala and Magdalena Rodríguez

213/2005 The effect of geographic expansion on the productivity of Spanish savings banks
Manuel Illueca, José M. Pastor and Emili Tortosa-Ausina

214/2005 Dynamic network interconnection under consumer switching costs
Ángel Luis López Rodriguez

215/2005 La influencia del entorno socioeconómico en la realización de estudios universitarios: una aproximación al caso español en la década de los noventa
Marta Rahona López

216/2005 The valuation of spanish ipos: efficiency analysis
Susana Álvarez Otero

217/2005 On the generation of a regular multi-input multi-output technology using parametric output distance functions
Sergio Perelman and Daniel Santín

218/2005 La gobernanza de los procesos parlamentarios: la organización industrial del congreso de los diputados en España
Gonzalo Caballero Miguez

219/2005 Determinants of bank market structure: Efficiency and political economy variables
Francisco González

220/2005 Agresividad de las órdenes introducidas en el mercado español: estrategias, determinantes y medidas de performance
David Abad Díaz
221/2005  Tendencia post-anuncio de resultados contables: evidencia para el mercado español
Carlos Forner Rodríguez, Joaquín Marhuenda Fructuoso y Sonia Sanabria García

222/2005  Human capital accumulation and geography: empirical evidence in the European Union
Jesús López-Rodríguez, J. Andrés Faña y Jose Lopez Rodríguez