Total factor productivity, intangible assets

and spatial dependence in the European regions

Barbara Dettori, Emanuela Marrocu, Raffaele Paci

Università di Cagliari and CRENoS

Abstract

In the last decade there has been an upsurge of studies on international comparisons of Total Factor Productivity (TFP). The empirical evidence suggests that countries and regions differ not only in traditional factor endowments (labour and physical capital) but mainly in productivity and technology. Therefore, a crucial issue is the analysis of the determinants of such differences in the efficiency levels across economies.

In this paper we try to assess these issues by pursuing a twofold aim. First, we derive a regression based measure of regional TFP which have the nice advantage of not imposing a priori restrictions on the inputs elasticities; this is done by estimating a Cobb-Douglas production function relationship for 199 European regions over the period 1985-2006, which includes the traditional inputs as well as a measure of spatial interdependences across regions. Secondly, we investigate the determinants of the TFP levels by analyzing the role played by intangible factors: human capital, social capital and technological capital. It turns out that a large part of TFP differences across the European regions are explained by the disparities in the endowments of such assets. This outcome indicates the importance of policy strategies which aim at increasing the level of knowledge and social capital as stressed by the Lisbon agenda. Estimation is carried out by applying the spatial 2SLS method and the SHAC estimator to account for both heteroskedasticity and spatial autocorrelation.

Keywords: Total factor productivity; human capital; social capital; technology; Europe.

JEL: C31, C33, O47, O52, R11

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. We would like to thank for their useful comments participants at 2008 ERSA Conference, 2009 SEA Conference and to DECA-CRENoS seminar. We have also benefited from fruitful discussions with Paola Zuddas. We thank Francesca Alberti, Giuliana Caruso and Marta Foddi for valuable assistance in preparing the database. We would also like to thank J.P. Elhorst for kindly making publicly available the matlab routines for estimating spatial models.

wp Crenos 2008/23

revised version July 2009

1. Introduction

Recent empirical literature on economic growth, both at country and regional level, has shown that the differences in the income levels are mainly due to disparities in the Total Factor Productivity (TFP) levels and to a lesser extent to the factors of production. Easterly and Levine (2001) report that more than 90% of the differences in growth rates among nations are explained by TFP rather than traditional factor accumulation. Moreover, a strong stylized fact that emerges from the empirical literature is that regional disparities are larger and more persistent when compared to cross countries differences, at least within the industrialized countries (see Magrini, 2004 for a review).

Since the differences in productivity turn out to depend on the efficiency levels, the attention of economists has been increasingly devoted to search for additional factors which may contribute to account for such disparities[1]. Several explanations for the TFP gap have been put forward, but among them a key role appears to be played by the intangible factors: human capital, social capital and technology. They create the base of the "knowledge economy" which, in turn, constitutes the most favourable environment to foster the economic performances of countries and regions, as stated by Lisbon declaration in 2000. As a matter of fact, in the industrialized economies the ability to compete in the open markets is increasingly based on production factors like the quality of labour, the degree of cohesion, the level of trust in the society and the accumulation of technological capital. However, there is a lack of systematic studies on the effects of different kinds of intangible assets on the economic performance at the regional level. In a number of studies human capital is often included as a determinant of the efficiency level, other works emphasize the effects of the knowledge-creation process and, only recently, social capital has been considered as a relevant variable in the context of explaining TFP variation across regions.

The main purpose and the novelty of this paper is to assess the effect of three different types of intangible assets on the economic performance at the regional level in Europe. Ideally, such a purpose would be nicely pursued by augmenting the traditional production function model with proxy variables for the intangible factors. However, for the European regions data on human, technological or social capital are not consistently available for all the regions over the entire sample period considered in this study. To deal with this severe lack of data we adopt the following two-step estimation strategy. First, we derive a measure of the Total Factor Productivity for the European regions by estimating a Cobb-Douglas production

[1] Since TFP is estimated using measured inputs, a possible cause of the disparities relies on measurement errors; moreover there may be problems of misspecification of the production function (Caselli, 2005).

function which includes only the traditional inputs, physical capital and labour. This is done in a panel data context - 199 regions over the period 1985-2006 – controlling for spatial dependence, time series non-stationarity and endogeneity. The estimated fixed effects represent an accurate measure of TFP at regional level which is directly derived from the production function estimation without imposing any (untested) restriction on the inputs elasticity parameters. In the second step we provide some interesting new evidence on the role played by intangible assets in determining the regional level of efficiency by including them as regressors in a model for the TFP data obtained in the first stage. It is worth stressing that, to the best of our knowledge, this is the first attempt to estimate "simultaneously" the effects of three different types of intangible capitals on the regional level of productivity.

The paper is organised as follows. In section 2 a detailed description of the data is presented; in section 3 we report and discuss the results for the Cobb-Douglas function estimation and for the derived TFP variable. In section 4 we present the main features of the intangible factors data followed by the discussion on the empirical evidence found on their effectiveness in enhancing regional productivity. Section 5 offers some concluding remarks.

2. Data descriptive analysis

2.1 Spatial patterns

The estimation of the production function relationship is based on a panel of 199 European regions observed over the period 1985-2006 (T=22); the regions belong to 15 member countries of the EU15 plus Switzerland and Norway. We follow the NUTS (Nomenclature des Unités Territoriales Statistiques) classification provided by Eurostat and select national and sub-national units, combination of NUTS 0, 1 and 2 levels, characterized by an adequate degree of administrative and economic control (see Appendix 1 for details).

Regional data on value added and labour units are obtained from the Cambridge Econometrics database. A detailed description of the variables used in this study, along with the indication of the sources, is presented in Appendix 2.

In what follows we discuss the geographical pattern of the variables included in the Cobb-Douglas function model. In order to reduce the degree of heterogeneity across regions all the series are rescaled with respect to the population size. In the map 1 the spatial pattern for value added (panel a), capital stock (panel b) and labour units (panel c) is depicted by reporting the quintile distribution of the time average of the per-capita values.

Map 1 shows evidence of significant cross-region dependence in the value added (panel a) distribution which follows a clear spatial scheme: among the worst performers are all the Greek and Portuguese regions, four Spanish regions (Extremadura, Andalucia, Galicia and Castilla-La Mancha) and the South of Italy (Calabria, Campania, Puglia, Sicilia, Basilicata and Sardegna). The top region is Denmark (63.6), at some distance from those that follow: Inner London (49.6), Zurich (44.2), Bruxelles (40.9) and Oslo (40.8). The map shows a partially different picture for high performing regions: all Swiss regions create a well defined cluster, as well as the Norwegian ones. A group of five German regions (Darmstadt, Karlsruhe, Stuttgart, Mittelfranken, Oberbayern, Salzburg) plus Salzburg in Austria make up another cluster. Also the southern areas of the United Kingdom form a cluster of high value added regions.

Since the series of the capital stock are not readily available from public databases, nor are they published by national statistical offices, the stock of physical capital K_{it} is calculated by applying the perpetual inventory method. For the sample 1985-2006 for each region i, in period t, K_{it} is constructed from the flow of gross investment in the previous period ($I_{i,t-1}$) and assuming an annual depreciation rate d equal to 10%, which is hypothesised to be constant over time and across regions:

$$K_{it} = (1 - d)K_{i,t-1} + I_{i,t-1} \tag{1}$$

The capital stock value for the initial year 1984 has been assumed equal to the cumulative sum of investment flows over the ten-years period 1975-1984.

Panel (b) of Map 1 shows the distribution of the physical capital stock (in per capita terms) across Europe: Central and Northern Europe show a large high-performance cluster, which starts from Steiermark in Austria, passes through most of the Southern German regions (Bayern and Baden-Württemberg) and ends with Denmark and southern regions of Norway. Detached from this cluster, one finds the capital regions of London (which shows the best performance) and Paris (Île de France). The regions displaying the worst performance are located at the European borders: in the West with Portugal and Spain, and in the South with the southern regions of Italy and the Greek ones.

As regards variability, value added shows a stronger dispersion of values, as indicated by a higher coefficient of variation (0.39 versus 0.31 for capital stock). As expected, data on units of labour (panel (c) Map 1) show lower variability, also confirmed by a lower variation coefficient (0.17) and a less accentuated geographical distribution of the centre–periphery type. Regions in the south of Germany always show the highest values, just as the Swiss ones, while the Norwegian regions do not appear in the first quintile. Note that the Scottish Highlands rank third with 0.64, after Inner London (0.67) and Bruxelles (0.87).

2.2 Testing for cross-section dependence

The presence of spatial dependence, evident in the maps discussed above, is also tested by means of the CD test proposed by Pesaran (2004) and the panel version of the Moran's *I* test (Kelejian-Prucha, 2001).

The CD test is a general test for general cross-section dependence which, has shown by Pesaran (2004), is applicable to a large variety of panel data models, including stationary and non-stationary dynamic heterogeneous panel with short T and large N, as is the case for the panel of data used in this study. The test is also robust to the presence of multi-breaks in slope coefficients and in the error variance. Correct size and satisfactory power are exhibited by the CD test even in small samples. The test, which is based on the average of the pair-wise correlation coefficients, is calculated as follows:

$$CD = \sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}\right)$$

where $\hat{\rho}_{i,j}$ are the sample estimates of the pair-wise correlation of the OLS residuals from individual regression in the panel; T=22, N=199. Following Pesaran (2004), in our case the residuals are obtained from models where the (log) of the variable being tested is regressed on a constant, a linear trend and on two of its own lags[2]. Under the null hypothesis of no cross-section dependence the test follows a standard normal distribution.

Although the CD test has power against spatial alternatives, we also compute the Moran's *I* test which is explicitly designed for such a case. The test, which under the null hypothesis is normally distributed, is calculated as:

$$I = \frac{\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\hat{u}_{it}\hat{u}_{jt}}{\left(T\sum_{i=2}^{N}\sum_{j=1}^{i-1}\left(w_{ij}+w_{ji}\right)^2\hat{\sigma}_i^2\hat{\sigma}_j^2\right)^{1/2}}$$

where $\hat{u}_{it}$ and $\hat{u}_{jt}$ are the residuals obtained from the same models estimated for the CD test, $\hat{\sigma}_i^2$ and $\hat{\sigma}_j^2$ are sample variances and w_{ij} are the elements of the weight matrix,

[2] Specifications with different dynamics and a model where the first difference of the variable is regressed on region-specific intercept (as done in Baltagi and Moscone, 2009) are also estimated yielding the same qualitative results.

capturing the spatial interconnections among regions, which in our case are measured by the inverse of the distance expressed in kilometres across regions.

The weight matrix W can be normalized in different ways. In most applied studies it is row-standardized, such that each row sum to unity; in this case the impact of all other regions on a particular region *i* is given by the weighted average of all regions' impacts[3]. Alternatively, the W matrix can be normalized with respect to a single normalization factor, its largest row sum or its largest characteristic root. In a recent paper Kelejian and Prucha (2009) argue that such a normalization is sufficient, while row-normalization imposes strong restrictions on the spatial process since each row of the W matrix is normalized in a different way.

In this study we apply the largest eigenvalue normalization, which, differently from the row-standardization, has the nice feature that the symmetry of the weights is preserved[4]; this is particularly important when W is an inverse distance matrix used to describe a "distance decay" type of economic behaviour, as stated in Anselin (1988) "scaling the rows so that the weights sum to one may result in a loss of that interpretation"[5].

The result for the CD and the Moran's *I* test are reported in Table 1. All the tests are highly significant leading to the rejection of the null hypothesis of no cross section (spatial) dependence among the European regions. To check the robustness of the results we calculate the Moran's *I* test allowing for different specifications of the *W* matrix, we considered both the largest eigenvalue normalization and the row-standardization for linear and square weights[6].

The CD test provides evidence that significant correlation is present between pairs of regions for all variables, while the Moran's *I* test suggests that such correlation is most likely due to spatial interdependence among regions. The estimation procedure presented in the next section will deal with this aspect of the data.

---

[3] In this case it is implicitly assumed that relative rather absolute distance matters; for a thorough discussion on normalization issues see Elhorst (2009).

[4] Note that, as emphasised by Anselin et al. (2008), the row-standardization has also the side effect that the sum of all the elements in *W* equals *N*, the number of cross-sectional observations, and that the induced asymmetry in the weights "is an unusual complication with significant computational consequences".

[5] See Baltagi et. al., 2008, for a discussion on the relevance of absolute distance *vs* relative distance in economic phenomena.

[6] Strong rejections (not reported in Table 1) are also found when the first and up to the second order contiguity matrix is considered.

2.3 Testing for non-stationarity

The possible non-stationarity property of the data is investigated by applying the CIPS test, recently proposed by Pesaran (2007). The test belongs to the so-called "second generation" of panel unit root tests and has the important advantage to overcome the main limitation of previous tests (see, among others, the widely applied tests suggested by Levin et al., 2002, Im et al., 1995, 2003 and Maddala and Wu, 1999), i.e. the assumption that the individual time series in the panel are cross-sectionally independently distributed; which is a questionable assumption, particularly in the context of cross-country (or region) regressions. The CIPS test, assuming a factor structure in the errors, deals with the cross-section dependence by augmenting the individual ADF regressions with the cross-section averages of the regressors and of the dependent variable. Consider wishing to test for presence of a unit root in the series y_t, of region i, the ADF regression is specified as follows:

$$\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{i,t-1} + \sum_{j=0}^{p} d_{ij} \Delta \bar{y}_{t-j} + \sum_{j=1}^{p} \delta_{ij} \Delta y_{i,t-j} + e_{it}$$

where the terms $\bar{y}_{t-1}$, $\Delta \bar{y}_{t-j}$ are the cross section averages for the lagged level and the lagged differences of y_t, respectively. The panel test is then calculated as the average of the individual t-test on the b_i coefficients. The test has satisfactory power and size even for relatively small panels; moreover, by means of an extensive Monte Carlo simulation study, Baltagi et. al. (2007) have shown that the CIPS test performs quite well when the cross-section dependence is originated by spatial correlation.

In this study we apply the truncated version of the test which limits the undue influence of extreme values that could occur when the time dimension is small; the test was calculated for both "intercept" and "intercept and trend" specifications and allowing for the lag order to be at maximum equal to 3 (p=0,1,2,3). The results are reported in Table 2; all the variables exhibit a non-stationary kind of behaviour with the exception of the labour variable, but only when p is selected to be equal to 0 or 1. On the contrary, the differenced series are stationary leading us to conclude that a panel unit root is present in the level series[7].

[7] Although not designed for the case of cross section dependence, we have also computed "first generation" tests (Levin-Lin-Chu test, Breitung t-stat, Im, Pesaran and Shin W-stat, ADF – Fisher, PP – Fisher, Hadri Z-stat) finding the same kind of results: the unit root hypothesis is marginally rejected only for the labour variable depending on the dynamic specification chosen.

3. Measuring total factor productivity

3.1 Estimation issues

As already mentioned in the introduction our two step strategy for the estimation of regional total factor productivity starts with the specification of the traditional Cobb-Douglas production function, which includes the conventional inputs, physical capital and labour, for a panel of 199 European region; it is formulated as:

$$Y_{it} = A_i K_{it}^{\beta_1} L_{it}^{\beta_2} \qquad (2)$$

with i=1, 2, …, N=199 and t=1,2, …, T=22 (sample period 1985-2006), where Y is value added at 2000 base prices; K is the stock of capital; L are labour units; A is the efficiency level. All variables are normalised to population in order to control for different size of the regions.

We first propose the estimated results of a sort of a "benchmark" model, which is a standard fixed effects model with time dummies of the log-linearized version of the Cobb-Douglas function reported above, we then propose different specifications of spatial panel models which take explicitly into account the geographical correlation among the European regions, as documented in the previous section.

Before presenting the estimation results in detail, given the non-stationary properties of the data, we discuss the evidence on cointegration tests carried out to check whether a long-run non-spurious relation exists among the variable included in model (2). We perform the well-known cointegration tests developed by Pedroni (1999, 2004) on the residuals obtained from the benchmark model; the tests are calculated for both the panel and group ADF and PP versions of the statistics and allowing for the two different specifications of the deterministic components, individual intercepts and individual intercepts and trends. The results reported in Table 3 allow to reject the null hypothesis of no cointegration in all cases considered. As the tests are derived under the assumption of cross-section independence, we also report the results for the series demeaned by subtracting the cross-section averages and for a model including the spatially lagged dependent variable (*WY*)[8], which explicitly accounts for the cross-section dependence. The evidence supports the existence of a long-run relationship among the variables included in the Cobb-Douglas production function model. Note that in this study, in the spirit of Pedroni (1999), we are interested "in the simple null hypothesis of no cointegration versus cointegration" in order to rule out any spurious

[8] In this case we are considering the spatially lagged variable as a variable which helps to explain the variation in the dependent one, rather than a simple left-hand side variable (Elhorst, 2009).

correlation among the variables, so we do not address the issue of cointegration vectors normalization; we are assuming that the particular normalization of the variables is the one represented by the production function relationship.

3.2 Econometric results

The estimation of spatial panel models is based on the following regression model:

$$y_{it} = \alpha_i + \beta_1 k_{it} + \beta_2 l_{it} + \delta W y_{it} + time\ \ dummies + u_{it} \qquad (3)$$

where low capital letters represents the log-transformed variables, a_i is the regional fixed effects, which, as will be discussed later on, represent our measure for total factor productivity, Wy_{it} is the spatially lagged dependent variable; we have also included time fixed effects to account for common shocks affecting the pooled regions. As already explained in section 2 the elements of matrix W are the spatial weights which are given by the square of the inverse of distance in kilometres; the matrix is then normalised by dividing each element by the largest characteristic root in order to maintain the symmetry of the distances. The choice to consider the square of the weights was driven by preliminary error diagnostics, the linear weights did not prove adequate to capture the spatial structure present in the data; the square values are supposed to be more informative and more powerful in discriminating between neighbouring and distant regions as they increase the relative weights of the closest ones.

Model (3) above is characterized by an "intrinsic" endogeneity problem arising from the inclusion of the spatial term, which induces a two-way causality in the neighbour relation in space ("each region is the neighbour of its neighbouring regions"). In this case consistent estimators are the ones derived from the maximum likelihood method or from the two-stage least squares (2SLS) one, based on the inclusion of instrumental variables. In the growing empirical literature on spatial models great care has been devoted so far in tackling the endogeneity due to the spatially lagged term while the potential endogeneity of the explanatory variables has often been overlooked, particularly in the panel data context[9]. In this study we attempt to take also into account the endogeneity between output and the production factors which can arise from system feedbacks or measurement errors. As the usual Durbin-Wu-Hausman (DWH) test points out that the stock of capital and (marginally)

[9] For cross-section analyses exception are represented by Kelejian and Prucha (2004, 2007), Anselin and Lozano-Gracia (2008), Fingleton and Le Gallo (2008) and Dall'Erba and Le Gallo (2008), see Elhorst (2007) for a panel application.

the labour units can be considered endogenous with respect to value added[10], we adopt the 2SLS estimation method in order to estimate the single structural equation we are interested in – the production function - without explicitly modelling the entire system relationships causing simultaneity (as in Fingleton and Le Gallo, 2008).

However, in the context of production function estimation there is clearly a paucity of adequate instruments, it is a very difficult task to find variables that, at the same time, are *directly* correlated with the explanatory variables but only *indirectly* correlated with the dependent variable so that they can be excluded from the regression without incurring in the omitted variable problem[11]. Following previous studies, in this work the instruments used for the productive factors are represented by their own values lagged up to two periods. Following Kelejian and Robinson (1993) and Kelejian and Prucha (1998), the spatially lagged term is instrumented by the explanatory variables lagged both in time and in space (pre-multiplied by the W matrix).

The estimation results are reported in Table 4A. The first column (*4.I*) reports the OLS estimation results for the "benchmark" model which, besides the individual intercepts and the dummy variables, includes only the traditional productive factors. At the bottom of the column we report the LM test for (remaining) spatial error correlation[12] and the Moran's I test; both tests indicate that, as expected, the estimated residuals are affected by spatial dependence. Regression 4.II shows the results of the spatial lag model estimated by applying the instrumental variable method, the estimated elasticities, 0.24 for capital and 0.29 for labour, are highly significant and the spatially lagged term ($\hat{\rho}$ =0.26) adequately captures the spatial dependence present in the data as shown by the insignificance of both diagnostic tests reported[13]. For all the 2SLS specification the Moran's I test is calculated as suggested in Anselin and Kelejian (1997) for the case of IV residuals.

To check the robustness of our results we also estimated model (3) by using a set of alternative instruments constructed by applying the 3-group method proposed by Kennedy

---

[10] Similar results are found when testing for weakly exogeneity of capital and labour within an error correction model framework; only labour can be considered weakly exogenous (the p-value for the null hypothesis that the adjustment term is zero in the labour ECM model is equal to 0.293).

[11] This point is also made by Temple (1999) for the case of growth regressions.

[12] The test for panel models, recently proposed by Anselin et al. (2008), is specified as follows: $LM_E = \frac{[e'(I_T \otimes W)e/(e'e/NT)]^2}{T\, tr(W^2 + W'W)}$, where e are the estimated errors and W is the weight matrix; under the null hypothesis the test is asymptotically distributed as a $\chi^2(1)$. Elhorst (2009) points out that its performance has still to be investigated when having panel data instead of cross-section data.

[13] We also calculate the panel version of the LM error test proposed by Anselin (1988) for testing for residual spatial autocorrelation in the presence of the spatially lagged dependent variable in cross-section models; the results do not change appreciably e do not affect the results.

(1992). For each explanatory variable the instrument takes the value -1, 0, or 1 according to whether the value of the instrumented explanatory variable is in the lower, middle or upper third of its ranking ranging from 1 to 199 in each period. Spatial lag of the 3-group instruments are considered for the spatially lagged dependent variable[14]. The results, reported in column 4.III, show that the estimated coefficients for the capital input and for the spatial lag term are higher with respect to those reported in the previous column, while the labour estimates is lower. Note also that model 4.III do not show evidence of residual spatial autocorrelation.

We also estimate a model which only accounts for the endogeneity of the spatial term (regression 4.IV), in this case by applying the ML method we found that the estimates are much more similar to those obtained for the 4.II regression.

Regression 4.V allows us to check for the robustness of a different measure of the labour input. We include the variable “hours worked per year” in place of “units of labour” to control for differences in the weekly worked hours provided for by different national legislation. The estimated coefficient (0.24) is quite similar to the one obtained in the 4.II specification[15].

Finally, as the estimation of the regional production function is relevant in its own right - beside serving as the base for measuring total factor productivity – we also investigate whether Objective 1 regions exhibit a significantly different performance with respect to the average of the regions; the results point out that, for the same level of capital and labour endowments, the Objective 1 regions show a considerable lower level of production; it is worth noting that in regression 4.VI no fixed effects are included and this results in higher estimated coefficients for both productive inputs while the spatially lagged term is associated with a very low coefficient; this seem to result in a misspecified model with spatially autocorrelated errors as diagnosed by Moran’s *I* test.

For all the estimated models discussed so far we guard against possible heteroskedasticity and remaining spatial correlation by applying the spatial heteroskedasticy and correlation consistent (SHAC) estimator for the variance-covariance matrix, proposed by Kelejian and Prucha (2007). The estimator is based on a set of assumption that is satisfied for

[14] Note that in a recent article, Fingleton and Le Gallo (2008) show that the 3-group instrument is a quasi-instrument as it can be completely uncorrelated with the error term, but in practical applications it yields low RMSEs.

[15] Note that the reported elasticities for the productive inputs are, in general, lower with those reported in previous studies (see, among others, Marrocu et al., 2001, for the Italian case and Ladu, 2006, for European regions); this is mainly due to the inclusion of both regional and time fixed effects, if they are excluded the estimated elasticities turn out to be 0.45 for the capital stock and 0.85 for labour, values which are comparable with those already reported in previous empirical works.

a large class of Cliff-Ord type models and is robust to measurement error in the spatial distance metric. Kelejian and Prucha (2007), by referring to a cross-section sample of n observations, assume that the error term, u, of a particular Cliff-Ord model with endogenous regressors (in our case we consider a panel spatial lag model, as the one reported in equation (3)), can be represented as u=Rε where ε is a vector of innovations and R is an nxn matrix of unknown elements; this formulation for the disturbance process allow for general unspecified form of correlation and heteroskedasticity. The asymptotic distribution of the IV estimator for the variance-covariance matrix is $\Psi = n^{-1}H'\Sigma H$, where H is the instruments matrix and $\Sigma(\sigma_{ij})$ is the variance-covariance matrix of u. Kelejian and Prucha (2007) show that the SHAC estimator for the (r,s)th element of $\hat{\Psi}$ is:

$$\hat{\psi}_{r,s} = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ir}h_{jz}\hat{u}_i\hat{u}_j K(d_{ij}/d_n)$$

where d_{ij} is the distance between unit i and unit j, while d_n is the bandwidth of a given kernel function (K) with the usual properties, K(0)=1, K(x)=K(-x) and K(x)=0 for |x|>1. Finally, Kelejian and Prucha show that small sample inference regarding the parameters vector, say δ, can be based on the approximation: $\hat{\delta} \sim N(\delta_0, n^{-1}\hat{\Phi})$,

where $\hat{\Phi} = n^2(\hat{Z}'\hat{Z})^{-1}Z'H(H'H)^{-1}\hat{\Psi}(H'H)^{-1}H'Z(\hat{Z}'\hat{Z})^{-1}$, Z is the regressors matrix (including both exogenous and endogenous variables) and $\hat{Z} = H(H'H)^{-1}H'Z$.

In the case of the models reported in table 4A we chose the Parzen kernel as defined in Andrews (1991)[16]. The bandwidth assumes the following values: 100, 300, 600 and 1200 kilometers; the first is a very short distance, the others distances correspond approximately to the lower decile, the lower quintile and the median of all the regional distances considered.

In table 4B we report the results for the t-ratios based on the SHAC estimates; in order to save space we confine the analysis to regression 4.II and 4.III of table 4A[17]. Overall the results obtained confirm the significance of all the regressors included in the model specifications considered; as expected t-ratios (standard-errors) tend to decrease (increase) as function of the bandwidth selected[18]. On the basis of the relative higher accuracy of the estimates, regression 4.II this is preferred to regression 4.III.

[16] The Parzen kernel, with x=d_{ij}/d_n, is defined as

$$K(x) = \begin{cases} 1-6x^2+6|x|^3 & \textit{for } 0 \le |x| \le 1/2 \\ 2(1-|x|)^3 & \textit{for } 1/2 \le |x| \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

[17] All the other results are available from the authors upon request.

[18] We also checked the robustness of our results with respect to the kernel function, similar results are obtained when using Bartlett weights (K(x)= 1-|x| for |x|≤1 and zero otherwise) instead of the Parzen ones.

In general the results reported in Table 4A offer further robust evidence on the relevance played by spillovers arriving from neighbouring regions in determining the production performance of the European regions and this, in turn, implies that a more rigorous representation of the spatial pattern present in the data cannot be further neglected.

3.3 Total Factor Productivity

From the fixed effects obtained from the estimation of regression 4.II we calculate the total factor productivity for each region which, as known, measures the efficiency in transforming physical capital and labour into output. The average values of TFP, computed as index relative to the European average, for the period 1986-2006 are reported in Map 1-panel (d). The best and worst ten regions are listed in Table 5.

Denmark is the leading region, with values nearly triple the European average, way ahead of the other regions in the ranking. Zurich, the capital regions of, Luxembourg, Belgium (Bruxelles) and Norway (Oslo) follow at some distance. Note that the efficiency index displays greater variability in the high end of the ranking, compared to the tail.

As for the geographical distribution of the index, we observe in the centre of Europe the concentration of high values around Switzerland and Western Germany regions. Moreover the TFP map shows the record levels of all Norwegian regions, of southern Ireland, North Eastern and Eastern Scotland, of a cluster of regions in the south area of UK, of three Dutch regions (Groningen, Utrecht and Noord-Holland) and the capital regions of France (Île de France), Sweden (Stockholm), Austria (Wien) and Italy (Lazio). Good results are also displayed by the Swedish regions, the French regions of Rhône-Alpes and Alsace, the western regions of Aquitaine and Midi-Pyrenees, and the centre-north of Italy (Trentino, Lombardia, Val d'Aosta and Emilia Romagna). Most of the regions of Portugal, Spain (except for the capital Madrid), Southern Italy and Greece (except for Sterea Ellada) stay in the lower part of the ranking. The lowest value is found, unexpectedly, for the region of Outer-London due to the presence of a high flow of labour commuting to Inner London.

The last panel of map 1 clearly depicts a spatial correlation pattern for the regional values of total factor productivity values across Europe; this is confirmed by the significant value (2.61, p-value 0.009) we found for the Moran's *I* test. In the following section we investigate the determinants of such spatial correlation within a spatial lag model framework.

4. The impact of intangible assets on TFP

4.1 Intangible assets

The purpose of this section is to assess the impact of the intangible assets on the level of TFP calculated in the previous section for the European regions. As stated in the introductory section, due to the lack of available long time series for variables such as social capital, our analysis, carried out in a cross-section framework, is confined to the year 2004.

More specifically we consider the effects of three types of intangible capitals: social capital, human capital and technological capital. In general, these intangible inputs are supposed to enhance the level of efficiency by creating a more favourable economic environment for firms; for this reason in the Lisbon agenda they are considered strategic in economic growth policies.

A complementary perspective based on micro data considers the intangible assets as part of business investment, like software, R&D expenditure, patents, economic competencies, employee training (OECD Secretariat, 1998). It is worth noting that Corrado et al. (2006) for the US firms estimate that total business investment in intangibles has roughly the same value of investment in tangible capital, therefore confirming the importance of including intangibles assets as determinants of productivity.

As mentioned in the introduction one of the novelty of this paper is to consider how productivity levels are influenced by social capital, which is an aspect often neglected in economic analyses as pointed out by Coleman (1990), Temple and Johnson (1998) and Tabellini (2008), among others. A high level of social capital in a certain area is often associated with widespread trust which, in turn, facilitates cooperation among the members of a community (Guiso et al. 2008), a reduction of transaction costs for both firms and consumers (Diani, 2004) and a wider diffusion of knowledge (Helliwell and Putnam, 1995). All these effects are proved to enhance the economic performance (Knack and Keefer, 1997). The literature provides several definitions of social capital (Glaeser et al., 2002); in general, it is considered as a set of informal norms and values, shared among members of a community, which allow them to cooperate. It is not an easy task to measure a complex phenomenon as social capital. In this paper, based on the broad definition given above, as a proxy for social capital we adopt the notion of social participation measured by the share of population that have taken part at least once in the last 12 months in social activities such as voluntary service, unions and cultural associations meetings over total population. The data at the regional level comes from the European Social Survey.

The distribution of social capital across the European regions for the year 2002 is presented in Map 2-panel (a)[19]. With reference to the geographical distribution of social capital, in Europe we see high value areas next to areas characterised by much lower values. The regions boasting the highest value of our indicator are located in the Scandinavian peninsula, in the four regions of Germany's Baden-Württemberg, in France's Mediterranean and Pyrenees areas and in the UK's South-West.

The literature has also emphasized the positive role of human capital on productivity level and growth (Mankiw et al., 1992; Benhabib and Spiegel, 1994). At the regional level a higher availability of well educated labour forces represents an advantage for the localization of innovative firms thus promoting local productivity (Rauch, 1993). As a proxy of human capital we use the share of population that has attained at least a university degree (ISCED 5-6) over total population.

The distribution of human capital across the European regions for the year 2002 is represented in the second panel of Map 2. Italy stands out for having all regions in the lowest class, while all other nations displaying values below the European average show greater variability and at least one region higher up in the rankings. This is the case with Portugal (with the Lisboa region) and Greece (with Attiki e Kentriki Makedonia). Note the excellent performance of Norway, Scotland, Finland's southernmost regions (Etela-Suomi and Lansi-Suomi) and eastern Spain (Cataluña, Aragona, Navarra, Pais Basco and Cantabria).

The inclusion in the production function of R&D expenditure as a direct measure of technology has been originally suggested by Griliches (1979) and afterwards the knowledge-capital model has been used in several contributions at firms level and also extended to macroeconomic models both at regional and country level. The idea is that technology is partly a public good, firms benefit from a higher degree of knowledge capital available in their areas since it leads to an increase in productivity. Recent contributions on the knowledge capital model include Madsen (2008) for the OECD countries; Fischer et al. (2008) for the European regions; Doraszerlsky and Jaumandreu (2008) for Spain. As an indicator for technological capital we use the number of patent applications adhering to the Patent Cooperation Treaty; the variable used in the estimation is calculated as the stock in the previous five years over total population[20]. The data have been regionalised on the basis of the

[19] For some regions in France, Germany and United Kingdom data are available at NUTS1 level so that we have assumed that value for the included NUTS2 regions. For a detailed description of the dataset see Parts (2008).

[20] We have also used R&D expenditure which is available for different years for each countries and the results are almost identical. The correlation coefficient between patents and R&D is equal to 0.82.

inventors' residence; in the case of patents with multiple inventors proportional quotas have been attributed to each region.

The distribution of technological capital across the European regions in the year 2002 is represented in the last panel of map 2; its per capita values show a large high-performance cluster, which starts from Rhône-Alpes (in France), passes through all Swiss regions and ends at the South-central part of Germany (Oberbayern, Freiburg, Stuttgart, Rheinhessen-Pfalz, Mittelfranken, Karlsruhe, Oberpfalz, Darmstadt, Tubingen, Unterfranken, Oberfranken). Close to this agglomeration are those of Düsseldorf and Köhln. These top performance regions are surrounded by other high performance countries. Detached from this cluster, one finds the capital region of Paris (Île de France). Sweden, Finland and Denmark show top-high innovation performance, suggesting the presence of a Scandinavian cluster. All southern European regions are characterised by very low levels of technological capital.

4.2 Econometric estimation and results

The effects of intangible assets are assessed by estimating the model specification reported below:

$$a_{it} = c + \beta_1 sk_{it-k} + \beta_2 hk_{it-k} + \beta_3 tk_{it-k} + \delta W a_{it} + \varepsilon_{it} \qquad (4)$$

where small letters indicate values in logs; *a* is the value of total factor productivity in each region in the year 2004 which was derived from regression 4.II in Table 4A; *sk* is social capital, *hk* is human capital and *tk* is technological capital; *i* are 199 regions. All variables are normalised to population in order to control for different size of the regions.

Due to potential system feedbacks and measurement errors, endogeneity problems can also be present in model (4); this are tackled by regressing the (log) level of TFP on the (log) level of the intangible asset for the year 2002 and, given that the dependent variable by construction do not exhibit time variability we also used the 3-group method instruments proposed by Kennedy (1992), as in the previous section. It is worth emphasising that in the case of TFP model the endogeneity issue is expected to be less problematic with respect to the case of the production function model as far as system feedbacks are concerned. From an economic standpoint, feedbacks between the productive inputs and the level of production are supposed to be direct (and stronger); on the other hand, when considering efficiency level the two-causality is reasonably weaker and the transmission channels from efficiency levels and intangible assets appear less clear. However, due to the inclusion of proxy variables, measurement errors remain a potential source of endogeneity.

The results for the TFP spatial lag model estimated by 2SLS are reported in Table 6A[21]. The first column presents the base model, all the intangible assets exhibit positive and significant coefficients: 0.14 for social capital, 0.16 for human capital and 0.07 for technological capital, thus confirming the crucial role played by this kind of productive factors. For the case of Italy this was documented also by Marrocu and Paci (2008) and by Di Giacinto and Nuzzo (2006), evidence of the positive effects of human capital on Italian regional productivity can be found in Di Liberto et al. (2008). In order to check for correct specification of the spatial pattern we calculate the IV-Moran test (Anselin and Kelejian, 1997), already mentioned in the previous section, which is specifically designed for the case of IV estimation. According to the test result no evidence of remaining residual spatial autocorrelation was found. Note that the coefficient of the spatially lagged term is strongly significant and high in value[22].

In the subsequent regressions 6.II-6.VI we try to assess which is the *crucial* distance to allow the benefits of one region to spill over the neighbouring ones. We calculate different weight matrices according to the distance selected; we start from a “short” distance of 0-300 km, the no zero links among the regions are therefore only those within such an interval in the unstandardized matrix[23]; we then consider three more distance ranges, each 300 km wide: 300-600, 600-900 and 900-1200[24]. Although we are aware that the wideness of the interval is completely arbitrary, on the basis of preliminary investigations we believe that we can derive some interesting insights on the spatial pattern of the regional spillovers.

The results for regression 6.II, where we include the four spatially lagged terms disaggregated according to the range distances, reveal that the relevant links are those within a 600 km distance. Only the first two spatially lagged terms are significant. This results is mainly driven by the fact that the weights of the W matrix are the inverse of the square distance, which penalizes interconnections between distant regions. Note also that this specification warrants a high significance to the main explanatory variables; moreover, with respect to the base model, the social capital coefficient increase in value from 0.14 to 0.18,

[21] The weight matrix used in the construction of the spatially lagged dependent variable is the same as the one adopted in section 3.

[22] Regression 6.I has also been estimated by including the Objective 1 dummy; the results point out that significant differences are present between the two groups of regions; moreover, the inclusion of the dummy variable makes the social capital variable irrelevant, while leaving all the other coefficients significant and of the same order of magnitude.

[23] The elements of the matrices used to construct the spatially lagged dependent variable are represented by the inverse of the square distance and each matrix is normalized with respect to its largest eigenvalue.

[24] As stated in the previous section 300, 600 and 1200 km correspond roughly to the lower decile, the lower quintile and the median of the regional distances.

while the estimated value of the other two intangible assets decreases only slightly. The Moran's *I* test does not signals residual spatial autocorrelation.

To check our results we then re-estimate regression 6.II by including only one spatially lagged term in turn. The evidence provided corroborates the previous finding, as only the 0-300 and the 0-600 lagged terms appear to be significant (regressions 6.III and 6.IV), however note that in the latter case and when considering successive distances (600-900 and 900-1200) the residuals are spatially correlated indicating that the exclusion of the links within the 300 km distance is detrimental for capturing the spatial dependency present in the data.[25]

Finally, we conduct a robustness check for the base specification by including a knowledge capital proxy instead of the technological capital one (regression 6.VII). The new proxy is calculated as the total funding by European Commission under the Fifth Framework Program (the program covers the 5-year period 1998-2002). Data on individual projects were regionalized by means of the address and postcodes of participants. In case of more than one participant, a proportional share of the funding was assigned to each of them. This new variable is expected to capture the effects of the creation of (new) knowledge on regional TFP; such effects are supposed to be more widespread and less specific, at least with respect to economic efficiency levels, than the ones induced by the patent activity. The coefficient of the knowledge capital variable is of the same order of magnitude as the one associated with technological capital, however its inclusion in the specification makes social capital more productive. This result may be due to possible complementarities between the two assets. A thorough investigation of such complementarities in enhancing efficiency levels is left for future research.

As for the case of panel models, we also calculate the SHAC estimates for the variance-covariance matrix of the empirical models reported in table 6A. Table 6B presents the t-ratios for the main specification, reg. 6.I; all the TFP determinants maintain their significance, thus confirming previous inference and the contribution of intangible assets in determining productivity[26].

[25] We have also considered three enlarging distance matrices for the spatial lag (0-600, 0-900, 0-1200): in all cases the results confirm the evidence provided by the base model.

[26] The same results are reached for all the specifications reported in table 6A (the only exception is technological capital which loses significance in model 6.IV when the bandwidth is set equal to 600 and 1200 km). All detailed results are available from the authors upon request.

5. Concluding remarks

The aim of this paper has been twofold. First, we have derived a regression based measure of regional TFP for Europe, which have the nice advantage of not imposing a priori restrictions on the inputs elasticities; this is done by estimating a Cobb-Douglas production function relationship which includes the traditional inputs as well as a measure of spatial interdependences across regions.

Secondly, we have investigated the determinants of the TFP levels by analyzing the role played by intangible factors: social capital, human capital and technological capital. This was motivated by a wide recent literature providing evidence which suggests that the economic performance across regions differ not only in traditional factor endowments (labour and physical capital) but mainly in technological, human and social capital.

The results for the production function model have confirmed the estimated elasticities already found in previous literature, but in our case these are obtained from an adequately specified model which properly accounts for the spatial pattern present in the data, without overlooking relevant econometric issues such as endogeneity and non-stationarity. The estimated TFP levels point out a concentration of high values around Switzerland and Western Germany, the highest values are found for all Norwegian regions, for southern Ireland, North Eastern Scotland and three Dutch regions; most of the Swedish regions, the French southern regions, the western regions of Aquitaine and Midi-Pyrenees and the centre-north of Italy display values above average, while most of the regions in Portugal, Spain and Greece are at the bottom of the ranking.

At the best of our knowledge, this paper represents the first attempt aimed at assessing the effects – at the same time - of three kinds of intangible assets on the regional efficiency levels in Europe. The estimated models have provided robust evidence on the role played by technological, human and social capital in enhancing economic growth and social cohesion. Moreover, the regional TFP levels are considerably affected by spatial spillovers which generate their strongest impacts in the range 0-600 km.

Our results, in turn, stress the importance of policy strategies aimed at accelerating the accumulation of this particular kind of endowments, as put forward in the Lisbon agenda.

References

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59, 817-858.

Anselin L (1988) *Spatial econometrics: methods and models*. Dordrecht: Kluwer.

Anselin L. and H.H. Kelejian (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors, *International Regional Science Review* 20, 153–182.

Anselin L. and N. Lozano-Gracia (2008) Errors in variables and spatial effects in hedonic house price models of ambient air quality, *Empirical Economics,* 34, 5–34.

Anselin L., J. Le Gallo and J. Hubert (2008), Spatial panel econometrics, in Mátyás L. and P. Sevestre (eds) *The Econometrics of Panel Data*. Berlin Heidelberg: Springer-Verlag.

Audretsch D. and M. Keilbach (2004) Entrepreneurship Capital and Economic Performance, *Regional Studies*, 38, 949-959.

Baltagi B., P. Egger and M. Pfaffermayr (2008) Estimating regional trade agreements effects on FDI in an interdependent world, *Journal of Econometrics*, 145, 194-208.

Baltagi B. and F. Moscone (2009) Health care expenditure and income in OECD reconsidered: evidence from panel data, University of Leicester, Department of Economics wp 09/05.

Baltagi, B. H., G. Bresson, and A. Pirotte (2007) Panel unit root tests and spatial dependence, *Journal of Applied Econometrics*, 22, 339-360.

Benhabib J. and M. Spiegel (1994) The role of Human Capital in Economic Development: evidence from aggregate cross-country data, *Journal of Monetary Economics*, 34, 143-174.

Caselli F. (2005) Accounting for Cross-Country Income Differences, in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Elsevier.

Coleman J. (1990) *Foundations of Social Theory*. Cambridge MA: Harvard University Press.

Corrado C.A., D.E. Sichel and C.R. Hulten (2006) Intangible Capital and Economic Growth, *FEDS wp* 2006-24.

Dall'erba S. and J. Le Gallo (2008) Regional convergence and the impact of European structural funds over 1989–1999: a spatial econometric analysis, *Papers in Regional Science,* 87, 219–244.

Di Giacinto V. and G. Nuzzo (2006) Explaining labour productivity differentials across Italian regions: the role of socio-economic structure and factor endowments, *Papers in Regional Science*, 85, 299-320.

Di Liberto A., F. Pigliaru and R. Mura (2008) How to measure the unobservable: a panel technique for the analysis of TFP convergence, *Oxford Economic Papers*, 60, 343-368.

Diani M. (2004) How associations matter. An empirical assessment of the social capital-trust-voluntary action link, in Prakash S. and Selle P. (eds) *Investigating Social Capital. Comparative perspectives on civil society, participation, and governance*. New Delhi: Sage.

Doraszerlsky U. and J. Jaumandreu (2008) R&D and Productivity: Estimating production functions when productivity is endogenous, *Harvard Institute of Economic Research*, wp 2147.

Easterly W. and R. Levine (2001) It's Not Factor Accumulation: Stylized Facts and Growth Models, *World Bank Economic Review*, 15, 177-219.

Elhorst J. P. (2009) Spatial Panel Data Models. In Fischer MM, Getis A (Eds.) *Handbook of Applied Spatial Analysis,* Ch. C.2. Springer: Berlin Heidelberg New York

Elhorst J.P., U. Blien and K. Wolf K (2007) New evidence on the wage curve: A spatial panel approach, *International Regional Science Review* , 30, 173-191.

Fingleton, B. and J. Le Gallo (2008) Estimanting spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties, *Papers in Regional Science*, 87, 319-339.

Fischer M., T. Scherngell and M. Reismann (2008) Knowledge spillovers and total factor productivity, *mimeo*, ERSA Conference, Liverpool.

Glaeser E., D. Laibson and B. Sacerdote (2002) An Economic Approach to Social Capital, *Economic Journal*, 112, 437-458.

Griliches Z. (1979) Issues in Assessing the Contribution of Research and Development to Productivity Growth, *The Bell Journal of Economics*, 10, 92-116.

Guiso L., P. Sapienza and L. Zingales (2008) Social Capital as Good Culture. Alfred Marshall Lecture*, Journal of the European Economic Association*, 6, 295-320.

Helliwell J. and R. Putnam (1995) Economic growth and social capital in Italy, *Eastern Economic Journal*, 21, 295-307.

Kelejian H.H. and I.R. Prucha (1998) A generalized spatial two stage least squares procedures for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of real estate finance and economics*, 17, 99-121.

Kelejian H.H. and I.R. Prucha (2001) On the asymptotic distribution of the Moran *I* test with applications, *Journal of Econometrics*, 104, 219-257.

Kelejian H.H. and I.R. Prucha (2004) Estimation of simultaneous systems of spatially interrelated cross sectional equations, *Journal of Econometrics*, 118, 27–50.

Kelejian H.H. and I.R. Prucha (2007) HAC estimation in a spatial framework, *Journal of Econometrics*, 140, 131–154.

Kelejian H.H. and I.R. Prucha (2009) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, forthcoming.

Kelejian H.H. and D.P. Robinson (1993) A suggested method of estimation fro spatial interdependent models with autocorrelated errors, and an application to a county expenditure model, *Papers in Regional Science*, 72, 297-312

Kennedy P. (1992) *A guide to econometrics*, Blackwell, Oxford.

Knack S. and P. Keefer (1997) Does social capital have an economic payoff? A cross-country investigation, *Quarterly Journal of Economics*, 112, 1251-1288.

Ladu M.G. (2006) Total Factor Productivity estimates: some evidence from European regions, *CRENoS wp* 2006/6.

Madsen J. (2008) Economic growth, TFP convergence, and the world export of ideas: a century of evidence, *Scandinavian Journal of Economics*, 110, 145-167.

Magrini S. (2004) Regional (di)convergence, in Henderson V. and J.F. Thisse (eds.) *Handbook of Regional and Urban Economics: Cities and Geography*. Amsterdam: Elsevier.

Mankiw N.G., Romer D. and D. Weil (1992) A contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107, 407-437.

Marrocu E. and R. Paci (2008) The effects of public capital on the productivity of the Italian regions, *Applied Economics*, forthcoming.

Marrocu E., R. Paci and R. Pala (2001) Estimation of total factor productivity for regions and sectors in Italy. A panel cointegration approach, *RISEC, Rivista Internazionale di Scienze Economiche e Commerciali*, 48, 533-558.

Moreno-Serrano R., R. Paci, and S. Usai (2005) Spatial spillovers and innovation activity in European Regions, *Environment and Planning A*, 37, 1793-1812.

OECD Secretariat (1998) *Measuring intangible investment: selected bibliography*, OECD, DSTI, Paris.

Paci R. and F. Pigliaru (2002) Technological Diffusion, Spatial Spillovers and Regional Convergence in Europe, in J. R. Cuadrado, M. Parellada (eds) *The European Monetary Union and Regional Convergence.* Springer, 273-292.

Parts E. (2008) Report on the new indicators of social capital in the European Union, *Working Paper IAREG* WP2/01.

Pedroni P. (1999) Critical values for cointegration tests in heterogeneous panels with multiple regressor, *Oxford Bulletin of Economics and Statistics*, 61, 653–670.

Pedroni P. (2004) Panel cointegration: asymptotic and .nite sample properties of pooled time series tests with an application to the PPP hypothesis, *Econometric Theory*, 3, 579–625.

Pesaran M.H. (2004) General diagnostic tests for cross section dependence in panels, University of Cambridge, *wp in Economics* 435 and CESifo *wp* 1229.

Pesaran M.H. (2006) Estimation and inference in large heterogenous panels with multifactor error structure, *Econometrica*, 74, 967-1012.

Pesaran M.H. (2007) A simple panel unit root test in the presence of cross section dependence, *Journal of Applied Econometrics*, 22, 265-312.

Rauch J. (1993) Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*. 34, 380-400.

Tabellini G. (2008) Institutions and Culture: Presidential Address, *Journal of the European Economic Association*, 6 255-294.

Tappeiner G., C. Hauser and J. Walde (2008) Regional knowledge spillovers: Fact or artifact? *Research Policy*, 37, 861–874.

Temple J. (1999) The new growth evidence, *Journal of Economic Literature*, 37, 112–156.

Temple J. and P.A. Johnson (1998) Social Capability and Economic Growth, *Quarterly Journal of Economics*, 113, 965-990.

Appendix 1. Regions and NUTS level

Country	**NUTS**	**Regions**
Austria	2	9
Belgium	1	3
Denmark	1	1
Finland	2	5
France (a)	2	22
Germany (b)	2	30
Greece	2	13
Ireland	2	2
Italy	2	20
Luxembourg	1	1
Netherlands	2	12
Norway	2	7
Portugal (a)	2	5
Spain (a)	2	17
Sweden	2	8
Switzerland	2	7
United Kingdom	2	37

(a) Territories outside Europe are not considered

(b) Berlin and East Germany regions are not considered

(c) Autonomous provinces of Trento and Bolzano are aggregated

Appendix 2. Data sources and variables description

Variable	Label	Source	Years	Measurement unit	Description
Value added	Y	Cambridge Econometrics	1985-2006	millions euros, 2000	
Capital stock	K	Own calculation	1985-2006	millions euros, 2000	
Units of labour	L	Cambridge Econometrics	1985-2006	thousands	
Hours worked	H	Cambridge Econometrics, own calculation	1985-2006	levels	total hours worked by employees per year
Population	POP	Cambridge Econometrics	1985-2006	thousands	
Human Capital	HK	Eurostat	2002, 2004	levels	people with a degree (ISCED 5-6)
Social capital	SK	European Social Survey Round 1 2002, Round 2 2004	2002, 2004	% of people over total population	% population that have taken part at least once in the last 12 months in social activities such as voluntary service, unions and cultural associations meetings
Technological capital	TK	OECD, REGPAT database	2000-2004	levels	patent applications at PCT (Patent Cooperation Treaty), stock for the previuos 5 years
Knowledge capital	KK	European Commission	1998-2002	euros, current	total funding by European Commission under the Fifth Framework Program (regionalized according to the research projects participants' address)
Dummy Objective 1 regions	DOb1	Eurostat	1985-2006		regions of the Objectives 1 program for the 2000-06 structural funds, including the transition regions

Table 1. Cross-section dependence tests

Test	Weight matrix	value added	capital stock	labour units
CD		119.88	83.38	115.27
		0.00	*0.00*	*0.00*
Moran's *I*	W	79.73	50.70	42.38
		0.00	*0.00*	*0.00*
	W-square weights	9.58	6.89	2.97
		0.00	*0.00*	*0.00*
	W-rstd	77.61	41.90	40.55
		0.00	*0.00*	*0.00*
	W-rstd -square weights	37.39	23.00	20.97
		0.00	*0.00*	*0.00*

All variables are in log-transformed per capita values

p -values are reported in italics;

W is the weight matrix normalized by dividing each element by the largest eigenvalue,

W-rstd is the same weight matrix row-standardized

Table 2. CIPS panel unit root tests

	Intercept only			Intercept and trend		
lags	value added	capital stock	labour units	value added	capital stock	labour units
		levels			*levels*	
p=0	-1.829	-1.256	-1.746	-2.174	-1.857	-2.641 **
p=1	-1.661	-1.687	-1.802	-1.943	-2.330	-2.833 **
p=2	-1.406	-1.621	-1.490	-1.480	-1.849	-2.391
p=3	-1.321	-1.735	-1.413	-1.328	-1.755	-2.453
		first differences				
p=0	-4.007 **	-2.508 **	-4.166 **			
p=1	-3.052 **	-2.380 **	-3.509 **			
p=2	-2.026 *	-1.918	-2.556 **			
p=3	-1.592	-1.705	-2.276 **			

Critical values are tabulated by Pesaran (2007) , Table II(a-c), we report the ones for T=20 and N=200 for the truncated version of the test:

Intercept case: -2.04 (5%); -1.99 (10%)

Intercept and trend case: -2.55 (5%); -2.49 (10%)

"**" and "*" indicates significance of the test at 5% and 10% level respectively

Table 3 Cointegration tests

Variables	Deterministic components	Pedroni tests	Statistic	P-value
Y, K, L	indiviudal intercepts	Panel PP-Statistic	-2.484	0.018
		Panel ADF-Statistic	-4.480	0.000
		Group PP-Statistic	-3.440	0.001
		Group ADF-Statistic	-9.638	0.000
Y, K, L	individual intercepts and trends	Panel PP-Statistic	-6.593	0.000
		Panel ADF-Statistic	-9.709	0.000
		Group PP-Statistic	-8.771	0.000
		Group ADF-Statistic	-14.638	0.000
Y*, K*, L*	indiviudal intercepts	Panel PP-Statistic	-6.615	0.000
		Panel ADF-Statistic	-8.896	0.000
		Group PP-Statistic	-5.498	0.000
		Group ADF-Statistic	-11.036	0.000
Y, K, L, WY	indiviudal intercepts	Panel PP-Statistic	-8.434	0.000
		Panel ADF-Statistic	-10.754	0.000
		Group PP-Statistic	-16.450	0.000
		Group ADF-Statistic	-16.333	0.000

Y, K and L stand for value added, capital stock and labour respectively; Y*, K*, L* are the same variables demeaned by subtracting the cross-section average

Null hypothesis: no cointegration

Alternative hypothesis: common autoregressive coefficient for panel specification or individual autoregressive coefficients for the group specification

Lag selection: Automatic SIC with a max lag of 4

Newey-West bandwidth selection with Bartlett kernel

Table 4A. Production function estimation with spatial lag model

Dependent variable: value added	**4.I**	**4.II**	**4.III**	**4.IV**	**4.V**	**4.VI**
Estimation method	*OLS*	*2SLS*	*2SLS*	*ML*	*2SLS*	*2SLS*
Instruments	--	*time and spatial lags*	*3-group instruments and their spatial lags*	--	*time and spatial lags*	*time and spatial lags*
Capital stock	0.270	0.240	0.306	0.270	0.279	0.444
	(21.3)	*(17.4)*	*(16.4)*	*(21.4)*	*(19.9)*	*(40.9)*
Units of labour	0.299	0.291	0.234	0.300		0.815
	(21.4)	*(17.3)*	*(10.1)*	*(21.6)*		*(40.9)*
Spatial lag		0.256	0.420	0.256	0.214	0.020
		(4.11)	*(2.9)*	*(2.7)*	*(3.6)*	*(2.7)*
Hours worked per year					0.242	
					(11.1)	
Dummy Objective 1 regions						-0.241
						(-29.2)
Regional fixed effect	yes	yes	yes	yes	yes	no
R^2 (pseudo)	0.98	0.98	0.98	0.98	0.98	0.81
N. obs	4378	3980	4378	4378	3980	3980
LM test for residual spatial correlation	45.251	0.169	0.043	0.790	2.045	1.865
p-value	*0.000*	*0.681*	*0.835*	*0.374*	*0.153*	*0.172*
Moran's *I* test on residuals*	6.279	0.340	0.183	1.530	0.167	1.897
p-value	*0.000*	*0.734*	*0.855*	*0.126*	*0.867*	*0.058*

Sample period: 1985-2006; all variables are normalised to population and log-transformed
Spatial weight matrix: square of the inverse of distance in km
Time fixed effects are included in all regressions
Aysmptotic t-statistic in parenthesis
R^2 (pseudo) is calculated as the ratio of the variance of the fitted values to the variance of the actual values
* For 2SLS the Moran's *I* test is calculated as the variant proposed in Anselin-Kelejian (1997) for IV residuals

Table 4B. Production function model SHAC estimates, Parzen kernel

Variable	Coefficients	*t*-ratios				
		Classical	SHAC			
			d_n=100	d_n=300	d_n=600	d_n=1200
Reg. 4.II (table 4A)						
Capital stock	0.240	17.366	11.132	9.227	7.708	6.589
Units of labour	0.291	17.329	9.617	8.038	6.840	6.225
Spatial lag	0.256	4.108	8.352	7.393	6.376	5.811
Reg. 4.III (Table 4A)						
Capital stock	0.306	9.951	7.768	6.452	5.493	4.837
Units of labour	0.234	10.066	7.579	6.350	5.683	5.498
Spatial lag	0.420	2.893	3.820	3.704	3.491	3.316

d_n is the kernel bandwidth in kilometers

Table 5. Total Factor productivity: best and worst regions

(Index: European average = 100; average 1986-2006)

Rank	DObj1	Region	Nation	TFP index
1	0	Denmark	Denmark	301
2	0	Zurich	Switzerland	200
3	0	Region Lemanique	Switzerland	176
4	0	Oslo og Akershus	Norway	175
5	0	Nordwestschweiz	Switzerland	174
6	0	Zentralschweiz	Switzerland	174
7	0	Luxembourg	Luxembourg	172
8	0	Bruxelles-Brussel	Belgium	163
9	0	Ostschweiz	Switzerland	159
10	0	Ticino	Switzerland	158
190	1	Extremadura	Spain	56
191	1	Alentejo	Spain	54
192	1	Ionia Nisia	Greece	53
193	1	Peloponnisos	Greece	51
194	1	Dytiki Ellada	Greece	51
195	1	Anatoliki Makedonia	Greece	50
196	1	Norte	Portugal	49
197	1	Centro (P)	Portugal	48
198	1	Ipeiros	Greece	47
199	0	Outer London	United Kingdom	40

Table 6A. Total Factor productivity and intangible assets

Dep. Variable: total factor productivity	6.I	6.II	6.III	6.IV	6.V	6.VI	6.VII
Social capital	0.137 *	0.185 ***	0.134 *	0.183 ***	0.178 **	0.180 **	0.332
	(1.678)	*(2.605)*	*(1.634)*	*(2.642)*	*(2.191)*	*(2.240)*	*(5.271)*
Human capital	0.165 ***	0.147 **	0.165 ***	0.160 ***	0.106 *	0.134 **	0.146
	(2.646)	*(2.412)*	*(2.641)*	*(2.713)*	*(1.771)*	*(2.064)*	*(2.324)*
Technological capital	0.067 ***	0.040 **	0.067 ***	0.030 *	0.061 ***	0.064 ***	
	(4.620)	*(2.466)*	*(4.639)*	*(1.850)*	*(3.575)*	*(4.320)*	
Knowledge capital							0.053
							(3.517)
Spatial lag dependent variable	0.741 ***		0.729 ***	0.990 ***	0.458	-0.072	0.780
	(2.903)		*(2.849)*	*(3.338)*	*(1.335)*	*(-0.173)*	*(3.127)*
Distances for spatial lag (in km)	all		0-300	300-600	600-900	900-1200	all
Spatial lag - distance 0-300 km		0.666 ***					
		(2.880)					
Spatial lag - distance 300-600 km		0.699 **					
		(2.275)					
Spatial lag - distance 600-900 km		0.248					
		(0.775)					
Spatial lag - distance 900-1200 km		0.002					
		(0.006)					
Square correl.	0.41	0.51	0.40	0.50	0.41	0.39	0.41
IV-Moran	-0.975	-0.915	-0.920	7.175 ***	6.377 ***	5.302 ***	-0.888
p-value IV-Moran	*0.329*	*0.360*	*0.358*	*0.000*	*0.000*	*0.000*	*0.375*

All regressions are estimated applying the 2SLS method, endogenous variables are instrumented with 3-group instruments (see section 4 for a detailed descript

Number of observations: 199

All variables are normalised to population and log-transformed

For human capital, social capital and technological capital the values refer to the 2002 year

Spatial weight matrix: square of the inverse of distance in km

All regression include a constant term

Aysmptotic t-statistic in parenthesis

Square correl. is the squared correlation between the predicted and actual values

IV-Moran is the Moran-I test proposed by Anselin and Kelejian (1997) for 2SLS residuals

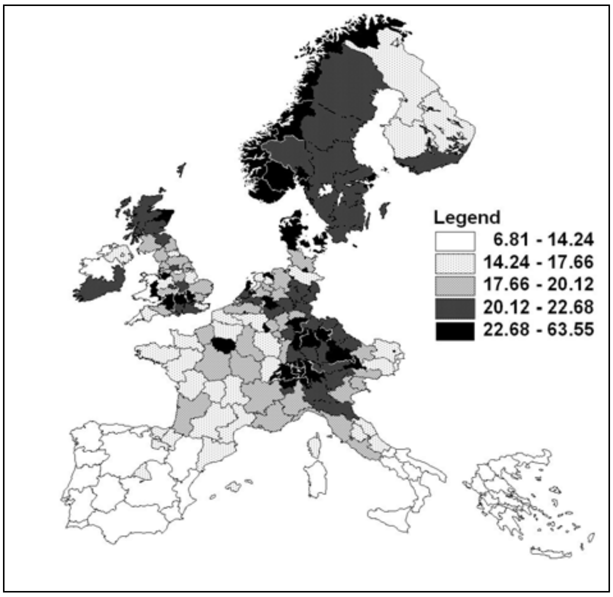
Level of significance: *** 1%, ** 5%, * 10%

Table 6B. TFP model SHAC estimates, Parzen kernel

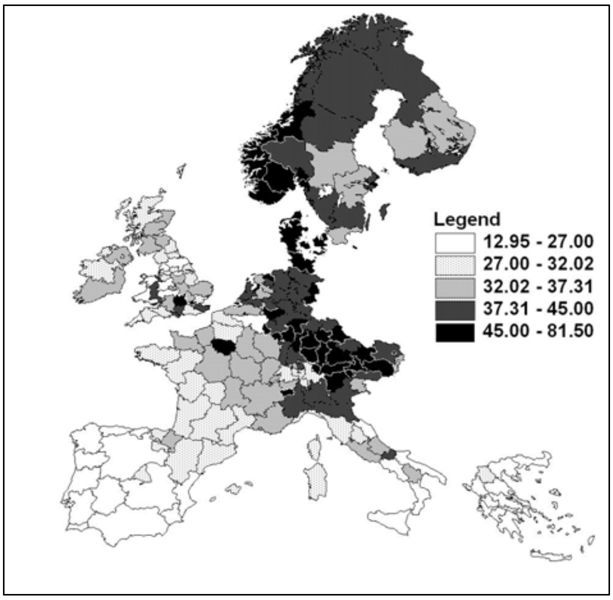
Variable	Coefficients	*t*-statistics				
		Classical	SHAC			
			d_n=100	d_n=300	d_n=600	d_n=1200
Social capital	0.137	1.678	2.090	2.044	1.870	1.824
Human capital	0.165	2.647	2.717	2.328	1.978	1.909
Technological capital	0.067	4.620	4.308	3.504	2.772	2.387
Spatial lag dependent variable	0.741	2.903	9.320	8.530	7.671	8.251

d_n is the kernel bandwidth

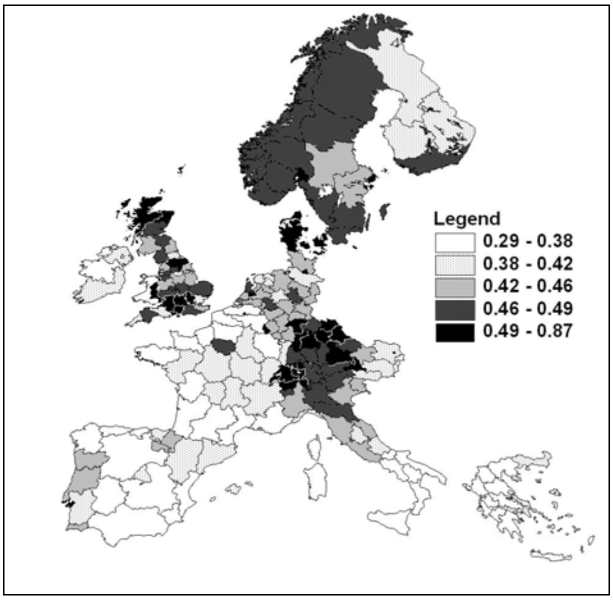
Map 1. Production function variables and estimated Total Factor Productivity



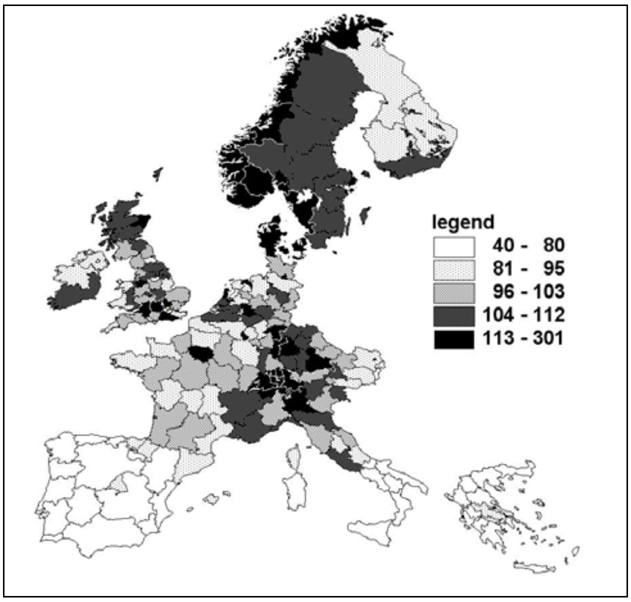
(a) Value added (per capita, thousands euro 2000)



(b) Capital stock (per capita, thousands euro 2000)



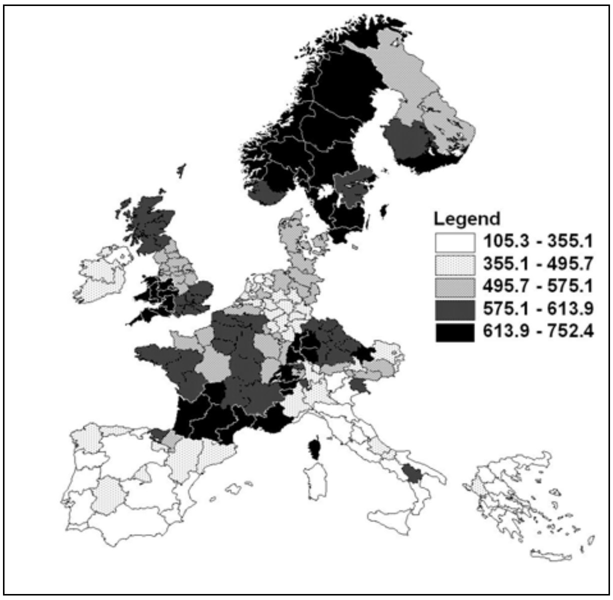
(c)Units of labour (per capita units)



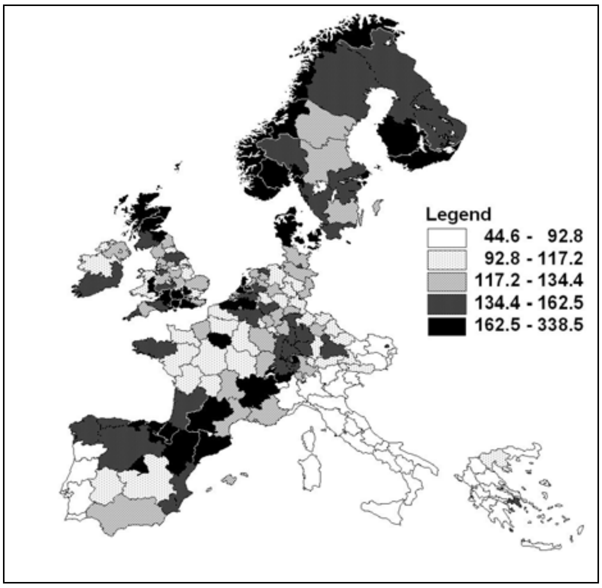
(d) Total factor productivity (index Europe = 100)

Values are average for the period 1986-2006

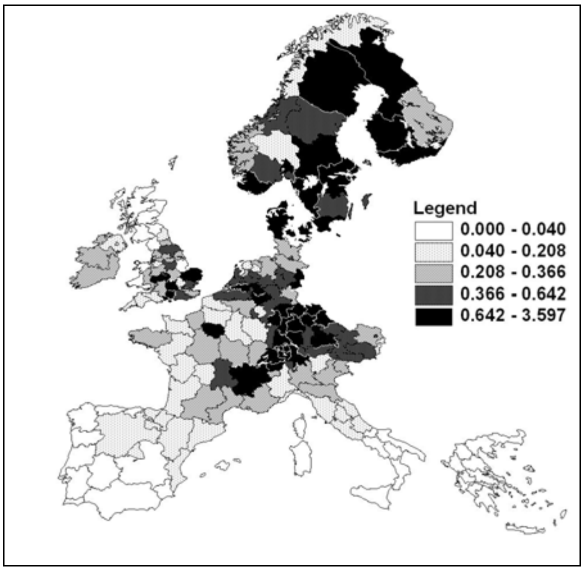
Map 2 Determinants of Total Factor Productivity: the intangible factors



(a) Social capital



(b) Human capital



(c) Technological capital

All values refer to the year 2002
(a) participation to social activities per thousands population
(b) inhabitants with a degree per thousands population
(c) patents PCT, 5-years stock, per thousands population