The effects of knowledge and innovation on regional growth: Nonparametric evidence

Marcos Sanso-Navarro[a,*], María Vera-Cabello[b]

[a]Departamento de Análisis Económico, Universidad de Zaragoza, Spain

[b]Centro Universitario de la Defensa de Zaragoza, Spain

November 2014

Abstract

This paper deals with the relationship between knowledge, innovation and regional growth. The study is carried out through the application of nonparametric estimation methods to European data at NUTS2 level. We provide evidence that the share of innovative firms plays a more relevant role in explaining regional growth than R&D expenditures. Further, inward FDI turns out to be a robust growth determinant. Our results also suggest that the effects induced by these variables are of a heterogeneous nature. As a byproduct of the analysis, we show that the estimation results from a local-linear kernel regression can be used for the identification of spatial patterns. In this respect, we find a cluster of innovation-driven labour productivity growth in Germany.

JEL classification: C14, C20, O18, R11

Key words: Regional growth, knowledge, innovation, nonparametric methods, nonlinearities

*Corresponding author. Address: Departamento de Análisis Económico. Facultad de Economía y Empresa. Gran Vía 2. 50005 Zaragoza (Spain). Tel: (+34) 876 554 629; Fax: (+34) 976 761 996; e-mail: marcossn@unizar.es.

1 Introduction

The 'Europe 2020' strategy for smart, sustainable and inclusive growth laid out by the European Union (EU) consists of the fulfilment of five main objectives at the end of the present decade (European Commission, 2010). Among these goals, it has been established that the share of expenditure in research and development (R&D) over gross domestic product (GDP) must be equal to 3% at the national level. This objective is primarily motivated by the endogenous growth literature as its fundamental premise is that deliberate decisions of rational agents can increase the productivity of labour which, in turn, generates economic growth. According to Romer (1990), decisions regarding R&D expenditures play a prominent role in these increments.

Given that there is a wide consensus on the importance of knowledge and innovation generated by R&D activities for regional development (Crespo Cuaresma et al., 2014), recent studies are more concerned with disentangling their differential effects (Capello and Lenzi, 2013; 2014). In line with empirical growth studies that take into account the possible presence of nonlinearities (Masanjala and Papageorgiou, 2004), there is also an interest in analyzing whether the influence that knowledge and innovation exert on growth is heterogeneous. Henderson et al. (2012a) propose the use of nonparametric estimation techniques to study the relevance and nonlinear influence of growth determinants. Nevertheless, carrying out this analysis in a regional context requires taking into account the possible presence of spatial dependence in the data (Basile, 2008). Given that the results obtained depend to a great extent on the way that this feature is modelled (Halleck Vega and Elhorst, 2013), McMillen (2012) advocate the use of nonparametric methods, which are flexible, to avoid this specification problem.

Although kernel regressions do not explicitly control for the spatial dependence across observations, their estimates can be consistent and asymptotically normal in the presence of this data feature (Robinson, 2011; Jenish, 2012). Further, Sanso-Navarro and Vera-Cabello (2014) provide evidence that the local-linear kernel estimator is more efficient than the alternative geographically-weighted regression method (GWR; Brunsdon et al., 1996). In the present paper, we propose the application of nonparametric estimation methods to study the relationship between, on the one hand, knowledge and innovation and, on the other, regional growth in the EU27 countries. Proceeding in this way, we will not only be able to determine if they are relevant for explaining growth, but also to analyze the possible presence of heterogeneity in the effects generated by these variables. Moreover, and as another contribution of the present paper, we show that the results obtained from the application of these estimators are useful for the identification of spatial patterns (Capello and Lenzi, 2013). This is possible by studying the geographical distribution of the estimated partial effects (gradients) using spatial analysis techniques (Anselin, 1995; Fischer and Getis, 2010).

The rest of the paper is structured as follows. Section 2 presents the empirical framework and the variables included in our analysis. Section 3 describes the nonparametric estimation methods on which the study is based. The relevance of knowledge and innovation as growth engines in EU regions, the heterogeneity of their effects and the presence of spatial patterns are assessed in Section 4. Section 5 concludes.

2 Empirical framework

Although numerous variables have been found to explain regional growth, there is a widespread consensus that knowledge and innovation are two robust growth engines (Cooke et al., 2011). In the present paper, we are not concerned with confirming this recurrent result in empirical growth studies. Instead, and following a recent trend in the literature, we try to disentangle the effects that knowledge and innovation separately exert on regional growth.

The reason for differentiating the effects generated by these two variables is that the efficiency gains derived from innovation activities depends on the strength of the local knowledge base (Bilbao-Osorio and Rodríguez-Pose, 2004; Rodríguez-Pose and Crescenzi, 2008). In addition, an interesting related question is to determine whether the influence of these two variables is characterized by the presence of spatial heterogeneity. For this reason, we also try to disentangle if knowledge and innovation have a nonlinear relationship with growth in EU regions.

[Insert Table 1 here]

With this aim, we adopt an empirical framework similar to that proposed by Capello and Lenzi (2013, 2014). This specification permits assessing the relevance of knowledge and innovation while controlling for other regional growth determinants. The analysis has been carried out with cross-sectional data for 262 NUTS2 regions[1] (EU27 countries).

[1] The choice of the areal unit of analysis is an important issue in empirical studies with aggregate spatial data sources. This is because different levels of aggregation can lead to different results, the so-called 'modifiable areal unit problem' (MAUP; Unwin, 1996). NUTS is the French acronym for *'Nomenclature of Territorial Units for Statistics'*, a hierarchical classification established by EUROSTAT to provide comparable regional breakdowns of EU member states. NUTS2 regions are defined

The variables included are described in Table 1, where their source, computation and sample period are also detailed. In particular, the empirical model corresponds to a growth regression that can be specified as follows:

$$g_i = \alpha + \beta KI_i + \gamma TE_i + \delta ED_i + \varepsilon_i; \quad i = 1, ..., n \tag{1}$$

where g_i denotes the average growth rate of real output in region i. KI_i is a vector containing proxies for the level of knowledge and innovation, TE_i is a vector reflecting territorial-enabling factors and ED_i includes control variables related to economic dynamism and socio-economic development. α is the intercept, ε is a zero-mean additive error and n is the number of regions.

As can be observed in Table 1, the dependent variable is the annual growth rate of real gross value added (GVA) per worker over the period 2005-2007, calculated with data from Cambridge Econometrics. The vector KI_i contains measures for the variables that play a central role in endogenous growth models: knowledge, human capital and innovation. The intensity of formal and basic knowledge is measured by the R&D expenditures as a share of GDP (R&D). The informal knowledge embedded in human capital is proxied by the share of managers and technicians over total employment (CAPABILITES). In line with recent studies, and because it is considered to have additional explanatory power for regional growth differences, the level of innovation is distinguished from R&D expenditures. Thus, a categorical variable reflecting the share of firms that introduce product and/or process innovations (INNOV) has also been included.

The effects of innovation and knowledge on regional growth cannot be analyzed without taking into consideration the social and institutional conditions. That is to say, the second group of variables included in vector TE_i try to reflect the influence of territorial-enabling factors for the effects of knowledge and innovation. In this regard, the infrastructures endowment has been measured by the rail and road potential accesibility over total area (INFRASTR). The functional specialization of a given region has been proxied by its corresponding share of blue collar occupations over total employment (FUNCTIONAL). The degree of entrepreneurship has been reflected by the share of self-employed over the total labour force, excluding the wholesale retail sectors (SELFEMPL). An indicator of the share of people trusting each other (TRUST) has also been introduced as a proxy for social capital.

according to a formal rather than a functional criteria, because they correspond to the level used for the implementation of regional policies. This institutional breakdown may influence the results, although to a lesser extent than if we were interested in modelling and analyzing regional spatial dependence.

The variables included in ED_i try to reflect the economic dynamism and stage of development of a given region. The employment growth rate (EMPL) and the location quotient of employment in knowledge-intensive services (KIS) try to measure the dynamics and level of specialization of the labour market, respectively. Further, the flow of inward foreign direct investment as a percentage of GDP (FDI) has also been considered in this third group of variables.

The influence of knowledge and innovation, as well as the rest of regional growth determinants, has been analysed through the application of nonparametric estimation methods. Although we are not unaware of the presence of spatial dependence between observations in the present context, these techniques have been applied because their estimates are consistent when this feature is present in the data (Robinson, 2011; Sanso-Navarro and Vera-Cabello, 2014). The following section is devoted to describing these methods, on which the empirical analysis is based.

3 Nonparametric kernel regression methods

To a great extent, the empirical analysis carried out in this study follows the approach proposed by Hall et al. (2007) and Henderson et al. (2012a). These studies exploited the fact that the relevance and nonlinear influence of the explanatory variables in nonparametric kernel regressions are revealed by their corresponding bandwidths when these parameters are determined using a least-squares cross-validation selection method. Moreover, and due to the flexibility of nonparametric estimation methods, it is not necessary to make any assumption about the functional form of the conditional mean or about the distribution of the error term.

The nonparametric counterpart of the empirical model in (1) can be expressed as:

$$g_i = m\left(X_i\right) + \epsilon_i; \quad i = 1, \ldots, n \tag{2}$$

where $X_i = (X_{i1}, X_{i2}, ..., X_{iq})$ is a vector of q variables related to regional growth - included in KI_i, TE_i and ED_i - and ϵ_i is the corresponding zero-mean additive error. Further, $m(\cdot)$ is the smooth unknown function for the conditional mean:

$$m(x) = E[g_i|X_i = x] \tag{3}$$

with $x = (x_1, x_2, ..., x_q)$ denoting the vector of growth determinants at which the conditional mean is evaluated.

One alternative for estimating the conditional mean function in (3) is by locally averaging the growth rates of the regions that are similar in terms of the values taken by their growth determinants. This method is known as the local-constant (or Nadaraya-Watson) kernel estimator:

$$\hat{m}\left(x\right) = \sum_{i=1}^{n} w_i g_i \tag{4}$$

Weights are non-negative, their sum is equal to one and they are given by

$$w_i = \frac{K(\frac{X_i - x}{h})}{\sum\limits_{j=1}^{n} K(\frac{X_j - x}{h})} \tag{5}$$

with

$$K(\frac{X_i - x}{h}) = k(\frac{X_{i1} - x_1}{h_1}) \cdot k(\frac{X_{i2} - x_2}{h_2}) \cdot ... \cdot k(\frac{X_{iq} - x_q}{h_q}) \tag{6}$$

and $k(\cdot)$ being a kernel function.

That is, the local-constant kernel estimator at x takes the average of the g_i values for the regions such that their X_i are in the neighborhood of x. The amount of information used to calculate the local average is determined by the bandwidths $h = (h_1, h_2, ..., h_q)$. A data-driven method for selecting these smoothing parameters is least-squares cross-validation, which consists of choosing h to minimize the following criterion:

$$CV(h) = \frac{1}{n}\sum_{i=1}^{n}(g_i - \hat{m}_{-i}(X_i))^2 M(X_i); \quad 0 \leq M(\cdot) \leq 1 \tag{7}$$

where $M(\cdot)$ is a weighting function and

$$\hat{m}_{-i}(X_i) = \sum_{l \neq i}^{n} \frac{g_l K(\frac{X_i - X_l}{h})}{\sum\limits_{l \neq i}^{n} K(\frac{X_i - X_l}{h})} \tag{8}$$

In other words, the criterion minimized by the cross-validation bandwidth selection is a trimmed version of the sum of squared residuals from a leave-one-out estimator of the conditional mean function. Following Li and Racine (2004), we have set $M(\cdot) = 1$.

Least-squares cross-validation bandwidth selection, in conjunction with the local-constant kernel estimator, is capable of automatically reducing the dimension of the problem when some of the regressors are irrelevant. More specifically, the irrelevant variables will be smoothed out as

$$k(\frac{X_{is} - x_s}{h_s}) \to k(0) \quad \text{when} \quad h_s \to \infty; \quad s = 1, 2, ..., q \tag{9}$$

Instead of the local-constant approximation, a linear regression through the regions with growth determinants in the same neighbourhood can be fitted. When a weighting function is included with this purpose, the estimation method is called the local-linear kernel regression. The aim is to estimate

$$g_i = a + b'(X_i - x) + e_i \tag{10}$$

As $(X_i - x)$ is used as the regressor, the intercept equals the conditional mean in (3). The estimation is based on solving the following optimization problem:

$$\min_{a,b} \sum_{i=1}^{n} (g_i - a - b'(X_i - x))^2 K(\frac{X_i - x}{h}) \tag{11}$$

It has been demonstrated that the solutions $\hat{a} = a(x)$ and $\hat{b} = b(x)$ are consistent estimators of the conditional mean function and of its partial derivative ($m^{(1)}(x) = \frac{\partial m(x)}{\partial x}$), respectively (Li and Racine, 2007). Due to its analogy to local least-squares, the local-linear estimation method nests the least-squares estimator as a special case for sufficiently large values of the bandwidth parameters. Moreover, the least-squares cross-validation method for bandwidth selection in the local-linear framework has the ability to select a large value of h_s when the conditional mean function is linear in x_s. On the contrary, it will select small values of the bandwidth parameter for regressors that have a nonlinear relationship with regional growth.

To sum up, the least-squares cross-validation bandwidth parameters for the local-constant regression will be used to draw conclusions regarding the relevance of regional growth determinants. The bandwidths for the local-linear estimation will allow us to determine its nonlinear influence. Given that the kernel function considered in the empirical analysis is the Gaussian one:

$$k(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}; \quad -\infty < v < \infty \tag{12}$$

we will conclude that a continuous growth determinant enters the conditional mean in an irrelevant fashion (local-constant regression) or linearly (local-linear) if its corresponding bandwidth parameter is greater than two times its sample standard deviation[2]. The versions of the estimation methods applied are those that allow us to handle both continuous and discrete variables in X_i. In this latter case, the upper bound is unity (Hall et al., 2007).

Before proceeding with the empirical analysis, it is worth noting that these estimators are based on the implicit assumption that each observation is independent and provides unique information. However, measurement problems, boundary mismatches or the presence of spillovers and externalities generate the presence of spatial autocorrelation among regions and, hence, implies a lack of independence. As pointed out by Rey and Janikas (2005), this dependence can result in misguided inferences and interpretations when using standard parametric estimation methods. Nevertheless, this is not necessarily the case for the local-constant and local-linear estimators. The conditions for their consistency and asymptotic normality when applied to spatially dependent data have been established by Robinson (2011) and Jenish (2012), respectively. Therefore, these properties can be added to the arguments in McMillen (2010) to advocate the use of nonparametric methods when dealing with spatial data.

4 Results

4.1 Growth determinants: Relevance and nonlinear effects

Our empirical analysis begins with the calculation of the bandwidth parameters with a least-squares cross-validation selection rule. Descriptive statistics for regional growth rates and each growth determinant included in the empirical model and their corresponding bandwidths are reported in Table 2.

It can be observed that the bandwidth parameters calculated for the local-constant estimation method are less than twice the sample standard deviation for most of the variables considered. The exceptions are R&D expenditures as a share of GDP, the proxy for the level of infrastructures and the employment growth rate. Therefore, the least-squares cross-validation bandwidth selection rule considers these variables as irrelevant for explaining labour productivity growth differences in EU regions during the period 2005-2007. These results show the importance of not only those variables related to endogenous growth models, but also their territorial-enabling factors and regional economic dynamism and development stage. Nevertheless, it can also be concluded that

[2] A performance evaluation of this procedure with relatively large numbers of relevant and irrelevant regressors in small samples can be found in Henderson et al. (2012a, pp. 148-152).

R&D expenditures are able to promote regional growth only when they are materialized in product and/or process innovation.

[Insert Table 2 here]

Having identified the relevant regional growth determinants, the next step in our analysis is to determine which of them exert a nonlinear influence. As has been explained in the previous section, this is related to the magnitude of the bandwidth parameter calculated by the least-squares cross-validation selection rule for the local-linear kernel regression estimator. The values obtained are reported in the last column of Table 2. They suggest that both the share of managers and technicians on total employment and the share of innovative firms exert a nonlinear influence on growth because their bandwidths are less than twice their sample standard deviation. With the exception of the social capital measure, this is also the case for the rest of control variables that are significantly related to growth.

Both the local-constant and the local-linear kernel estimators assume that the observations are independent and, hence, do not explicitly account for the presence of spatial dependence when applied in the present context. In order to analyze the extent to which these methods and the empirical specification considered in our analysis capture this feature of European regional data, the global Moran's I test has been calculated for the residuals of the kernel regressions using two k-nearest ($k = 5, 10$) neighbours matrices[3].

The resulting test statistics, along with their p-values, are reported in the lower panel of Table 2. The null hypothesis of the global Moran's I test is the absence of spatial autocorrelation. It cannot be rejected at the 5% significance level either for the local-linear or the local-constant estimations. This can be interpreted as evidence that kernel regressions are able to control for the spatial dependency in the data when explanatory variables are close not only in the variable space but also in the geographical space, as is the case in our data. As expected, the location quotient in KIS sectors is the only variable for which the null hypothesis of no spatial autocorrelation cannot be rejected[4].

[3] A distance-based weights matrix has not been used because the Canary Islands are included in our sample. The minimum distance to consider in this case for all the regions to have, at least, one neighbour is very high.

[4] These results are available from the authors upon request.

4.2 Identifying territorial patterns

A common practice to obtain partial slopes in multivariate settings is to select an explanatory variable and hold the remaining covariates at specific values (like their sample means). Nevertheless, kernel regressions can be used to calculate the marginal effects (gradients) of a covariate at a given point. They are obtained as the derivative of the conditional mean in (3) at the value x. Hence, the marginal effect of a covariate for each observation is calculated at the observed values of all the covariates for this same observation.

In this line, Henderson et al. (2012b) propose 2-dimensional figures (45^o plots) that help to clarify the heterogeneity that stems from the estimates of multivariate models. The corresponding plots for the statistically significant gradients for the six covariates that, according to the results in the previous subsection, have a nonlinear relationship with EU regional growth are displayed in Figure 1. In addition, the mean value and relevant quartiles for all these gradients (significant and non-significant) are reported in Table 3.

[Insert Figure 1 here]

The heterogeneous character of the influence of these growth determinants is confirmed by the six 45^o plots. In particular, the share of managers and technicians and, in line with the results in Capello and Lenzi (2013, 2014), of blue collar occupations on total employment tend to exert a negative influence on growth. This result may be reflecting convergence issues not accounted for by the empirical framework considered. The reason is that the initial level of productivity is not controlled for and may be related to the dates that these variables refer to. In addition, the share of innovative firms seems to have a negative relationship with regional growth. However, this result is a consequence of the high standard errors of the estimated partial effects for this variable, what may be related to its discrete nature. As can be observed in Table 3, when all the gradients of this proxy for innovation are taken into account, both its mean and its median and upper quartiles are positive. Finally, the estimated partial effects suggest that specialization in KIS and, to a greater extent, inflows of FDI have growth-enhancing effects.

[Insert Table 3 here]

Following the related literature (Funke and Niebuhr, 2005), the heterogeneity found in the partial effects of these growth determinants may be driven by the presence of threshold effects. The extent to which the variables related to knowledge and innovation generate this type of nonlinearity has been analyzed by comparing the kernel density functions of their significant partial effects, depending on whether they are above or below the sample median. This comparison is plotted in Figure 2. Each column refers to the variable that generates the threshold effects, that is, the variable that takes values above or below the European sample median. Each row refers to the variable that experiences the threshold effect and, thus, for which the densities of the gradients are compared. In addition, a formal comparison has been carried out by applying the test of equal density functions proposed by Li et al. (2009), that is also based on the least-squares cross-validation bandwidth selection. The test statistics obtained, along with their corresponding bootstrap p-values (399 replications), are also reported in each graph.

[Insert Figure 2 here]

According to this test, the share of managers and technicians on total employment is the growth determinant related to knowledge and innovation that tends to be affected by threshold effects. They are generated both by this variable and the share of innovative firms. It can also be observed that there are a higher number of non-significant partial effects of the knowledge embedded in human capital in the regions where this variable is above the EU median. In addition, regions with a lower endowment of human capital tend to obtain fewer benefits from it. However, this variable tends to exert a more positive influence on growth in regions with a lower share of innovative firms. The latter also have a much higher frequency of negligible effects generated by innovation. Therefore, it can be stated that innovation results have a positive influence on growth once a threshold value has been achieved.

The GWR estimation method provides intercept and slope parameters for each region in the sample by running a sequence of local-linear regressions using subsets of data that are close in the geographical space, instead of in the variable space. As pointed out by McMillen (2010), GWR is a special case of standard non-parametric regression procedures that has attracted the attention of researchers, who have neglected the advantages of other estimators. For this reason, we complete our analysis by showing that the estimated gradients from the local-linear kernel estimator allow us to identify spatial patterns. This has been done by constructing cluster maps with the local indicator of spatial association (LISA; Anselin, 1995) for these partial effects.

[**Insert Figures 3 and 4 here**]

The LISA cluster maps[5] for the partial effects of the share of innovative firms and the inward flows of FDI are shown in Figures 3 and 4, respectively. The former suggests that there is a significant 'high-high' spatial correlation in the effects of the share of innovative firms in German regions. In line with Capello and Lenzi (2013), this implies that there is not only a high degree of innovation in the 'European Science-based area' but also that these regions are where innovation has a higher positive influence on growth. Further, there are two clusters of 'low-low' spatial association in the effects of innovation in Italian and Spanish regions. Figure 4 shows the LISA cluster map for the gradients of inward FDI flows in a given region. Although this variable turns out to be a robust driver of growth, French and Italian regions are those that obtain a higher benefit from inward FDI.

5 Concluding remarks

This paper has applied nonparametric kernel estimation methods to study the relationship between knowledge, innovation and growth in European regions. We find that the share of innovative firms explains labour productivity growth differences at a NUTS2 level. However, our results suggest that R&D activities lose their relevance when jointly considered with innovation and the knowledge embedded in human capital. We also obtain evidence regarding the important role of inward FDI flows as a growth determinant. In line with related studies, we have found the presence of a nonlinear relationship between regional growth and its determinants. The heterogeneity of the effects that innovation exert on growth has been confirmed by the partial effects obtained from a local-linear kernel estimator. As a novelty, we have shown that these gradients can be useful in detecting spatial patterns.

[5]LISA cluster maps have been constructed for a k-neighbours weights matrix with $k = 10$. Considering a smaller number of neighbouring regions leads to similar conclusions.

Our findings suggest that EU policies should take into account not only that regions have different characteristics but also that these policies affect growth in different ways. In addition, a policy based on the establishment of a target for the level of R&D expenditures seems not to be appropriate at a regional level. It would be much more important to intervene in order to ensure that these activities really contribute to knowledge accumulation through innovation results. Policies should be devoted to promoting activities intended to attract FDI.

Acknowledgements

The authors are grateful to Christopher Parmeter for his help with programming and Fernando Gutierrez for providing access to the fDi Intelligence database. Financial support from Ministerio de Ciencia e Innovación (ECO2011-22650 project) and the Regional Government of Aragón (ADETRE Research Group) is also acknowledged.

References

[1] Anselin, L., 1995. Local indicators of spatial association – LISA. Geographical Analysis 27(2), 93-115.

[2] Basile, R., 2008. Regional economic growth in Europe: A semi-parametric spatial dependence approach. Papers in Regional Science 87(4), 527-545.

[3] Bilbao-Osorio, B. and A. Rodríguez-Pose, 2004. From R&D to innovation and economic growth in the EU. Growth and Change 35(4), 434-455.

[4] Brunsdon, C., A. S. Fotheringham and M. Charlton, 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. Geographical Analysis 28(4), 281-298.

[5] Capello, R. and C. Lenzi, 2013. Territorial patterns of innovation and economic growth in European Regions. Growth and Change 44(2), 195-227.

[6] Capello, R. and C. Lenzi, 2014. Spatial heterogeneity in knowledge, innovation and economic growth nexus: Conceptual reflections and empirical evidence. Journal of Regional Science 54(2), 186-214.

[7] Cooke, P., B. Asheim, R. A. Boschma, R. Martin, D. Schwartz and F. Tödtling, 2011. Handbook of Regional Innovation and Growth. Chentelham: Edward Elgar.

[8] Crespo Cuaresma, J., G. Doppelhofer and M. Feldkircher, 2014. The determinants of economic growth in European regions. Regional Studies 48(1), 44-67.

[9] European Commission, 2010. Europe 2020: A strategy for smart, sustainable and inclusive growth. Brussels, March.

[10] Fischer, M. M. and A. Getis, 2010. Handbook of applied statistical analysis: Software tools, methods and applications. Berlin: Springer-Verlag.

[11] Funke, M. and A. Niebuhr, 2005. Threshold effects and regional economic growth – evidence from West Germany. Economic Modelling 22(1), 61-80.

[12] Hall, P., Q. Li and J. S. Racine, 2007. Nonparametric estimation of regression functions in the presence of irrelevant regressors. Review of Economics and Statistics 89(4), 784-789.

[13] Halleck Vega, S. and J. P. Elhorst, 2013. On spatial econometric models, spillover effects, and W. University of Groningen, Mimeo.

[14] Henderson, D. J., C. Papageorgiou and C. F. Parmeter, 2012a. Growth empirics without parameters. Economic Journal 122, 125-154.

[15] Henderson, D. J., S. C. Kumbhakar and C. F. Parmeter, 2012b. A simple method to visualize results in nonlinear regression models. Economics Letters 117(3), 578-581.

[16] Jenish, N., 2012. Nonparametric spatial regression under near-epoch dependence. Journal of Econometrics 167(1), 224-239.

[17] Li, Q., E. Maasoumi and J. S. Racine, 2009. A nonparametric test for equality of distributions with mixed categorical and continuous data. Journal of Econometrics 148(2), 186-200.

[18] Li, Q. and J. S. Racine, 2004. Cross-validated local linear non-parametric regression. Statistica Sinica 14, 485-512.

[19] Li, Q. and J. S. Racine, 2007. Nonparametric econometrics. Theory and practice. Princeton University Press, New Jersey.

[20] Masanjala, W. H. and C. Papageorgiou, 2004. The Solow model with CES technology: Nonlinearities and parameter heterogeneity. Journal of Applied Econometrics 19(2), 171-201.

[21] McMillen, D. P., 2010. Issues in spatial data analysis. Journal of Regional Science 50(1), 119-141.

[22] McMillen, D. P., 2012. Perspectives on spatial econometrics: Linear smoothing with structured models. Journal of Regional Science 52(2), 192-209.

[23] Rey, S. and M. Janikas, 2005. Regional convergence, inequality and space. Journal of Economic Geography 5(2), 155–176.

[24] Robinson, P. M., 2011. Asymptotic theory for nonparametric regression with spatial data. Journal of Econometrics 165(1), 5-19.

[25] Rodríguez-Pose, A. and R. Crescenzi, 2008. Research and development, spillovers, innovation, systems, and the genesis of regional growth in Europe. Regional Studies 42(1), 51-67.

[26] Sanso-Navarro, M. and M. Vera-Cabello, 2014. Non-linearities in regional growth: A non-parametric approach. Papers in Regional Science, forthcoming. DOI: 10.1111/pirs.12112.

[27] Unwin, D. J., 1996. GIS, spatial analysis and spatial statistics. Progress in Human Geography 20(4), 540-441.

Table 1. Variable description, construction and data sources.

Variable	Indicator	Measure	Computation	Years	Source
GROWTH	Real gross value added (GVA) per worker	Growth	Annual rate of growth	2005-2007	Cambridge Econometrics
CAPABILITIES	Knowledge embedded in human capital	Share of managers and technicians	Share of ISCO codes 13 and 31 over total employment	Average 1997-2001	European Labour Force Survey
INNOV	Innovation of product and/or process	Firms introducing innovations	Share of firms. Ranked values (0-8)	2002-2004	ESPON
R&D	Research and development (R&D)	R&D expenditures	Share of R&D expenditures over GDP	Average 2000-2002	EUROSTAT
INFRASTR	Infrastructure endowment	Rail and road potential accesibility over total area	Rail and road potential accesibility over total area	2001	ESPON
FUNCTIONAL	Functional specialization	Share of blue collar occupations	Share of ISCO codes 7 and 8 over total employment	Average 1997-2001	European Labour Force Survey
SELFEMPL	Self-employment	Share of self-employment	Share of self-employment over total labour force (excluding wholesale retail sectors)	Average 1999-2004	European Labour Force Survey
TRUST	Social capital	Trust	Share of people trusting each other	1999-2000	European Value Survey
EMPL	Employment growth rate in energy and manufacturing	Employment dynamics	Annual rate of growth	2002-2004	Cambridge Econometrics
FDI	Foreign direct investment (FDI)	Inward FDI flows	Inward FDI flows as percentage of GDP	Average 2003-2004	fDi Intelligence and EUROSTAT
KIS	Specialization in knowledge intensive services (KIS)	Location quotient in KIS sectors	Location quotient of employment in KIS sectors	2002-2004	EUROSTAT

Table 2. Descriptive statistics for regional growth and its determinants and least-squares cross-validation bandwidths.

	Mean	Median	SD	Min	Max	Local-constant		Local-linear	
GROWTH	2.69	2.40	2.20	-2.55	10.32	—		—	
Knowledge and innovation									
CAPABILITIES	7.30	7.15	2.54	2.96	17.94	1.26		1.14	
INNOV	4.44	4	1.88	1	8	0.51		0.41	
R&D	1.40	1.01	1.21	0.07	8.85	3.00*		—	
Territorial-enabling factors									
INFRASTR	35.52	8.88	63.64	0.02	562.89	584.70*		—	
FUNCTIONAL	23.93	23.74	6.35	7.86	39.50	3.30		4.95	
SELFEMPL	12.23	10.21	6.43	3.45	38.08	2.19		10.53	
TRUST	30.97	28.09	15.65	0	82.35	6.35		34.35**	
Economic dynamism and development stage									
EMPL	-1.68	-1.79	2.91	-11.65	6.87	21.63*		—	
FDI	2.31	0.57	5.07	0	41.81	3.03		8.75	
KIS	0.96	0.94	0.17	0.51	1.97	0.06		0.15	
					R^2	0.97		0.89	
					I(k=5)	-0.02	*(0.36)*	0.02	*(0.26)*
					I(k=10)	-0.03	*(0.15)*	0.03	*(0.08)*

Note: * denotes that the variable is smoothed out of the regression and ** indicates that the variable enters linearly. I(·) is Moran's *I* test statistic for a k-nearest neighbours specification of the weights matrix. p-values in parentheses.

Table 3. Partial effects for continuous and relevant regional growth determinants.

	Mean	Q1	Q2	Q3
CAPABILITIES	-0.09	-0.37	-0.11	0.12
	(0.07)	(0.19)	(0.27)	(0.09)
INNOV	0.04	-0.27	0.04	0.38
	(0.06)	(0.19)	(0.06)	(0.58)
FUNCTIONAL	-0.01	-0.12	-0.04	0.06
	(0.02)	(0.04)	(0.03)	(0.07)
SELFEMPL	0.01	-0.08	0.02	0.09
	(0.02)	(0.02)	(0.02)	(0.02)
TRUST	-0.02	-0.03	-0.01	2.92E-03
	(0.01)	(0.01)	(0.01)	(0.01)
FDI	0.18	0.03	0.17	0.38
	(0.01)	(0.14)	(0.15)	(0.07)
KIS	0.48	-3.27	-1.46	2.85
	(0.70)	(1.30)	(2.54)	(1.04)

Note: Partial effects are the estimated derivatives from the local-linear nonparametric regression. Bootstrap standard errors in parentheses (399 replications).

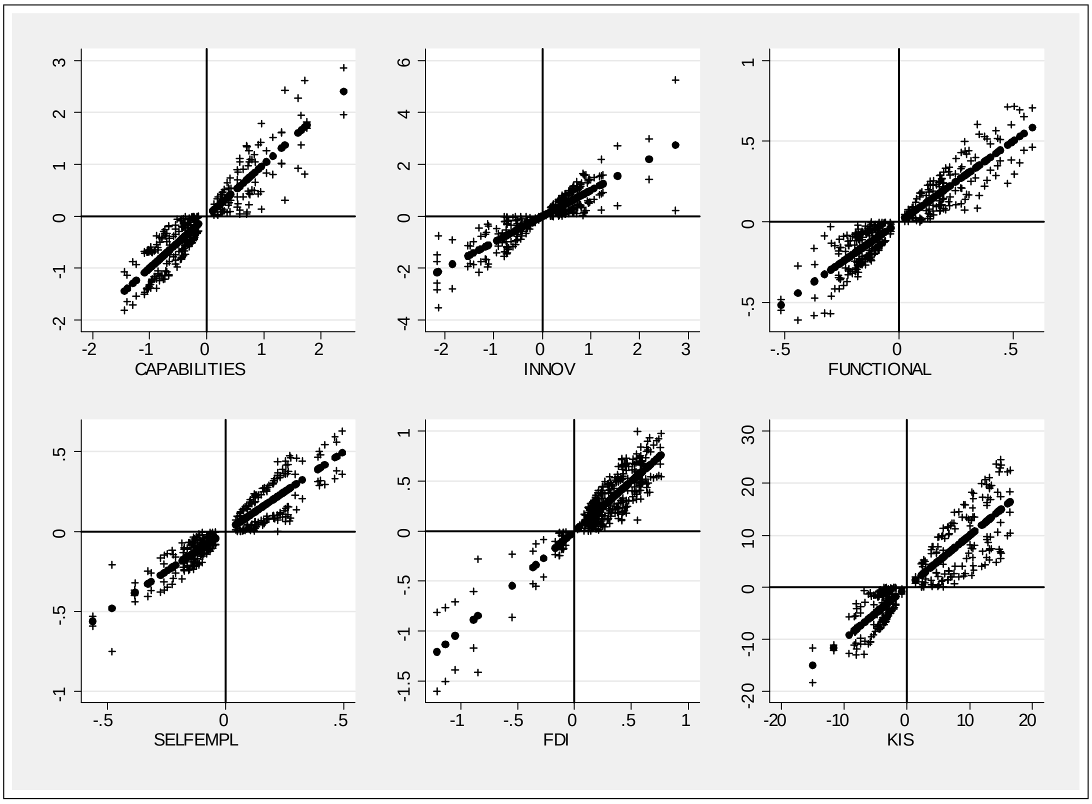

Figure 1: 45° plot of the statistically significant estimated gradients for selected regional growth determinants.

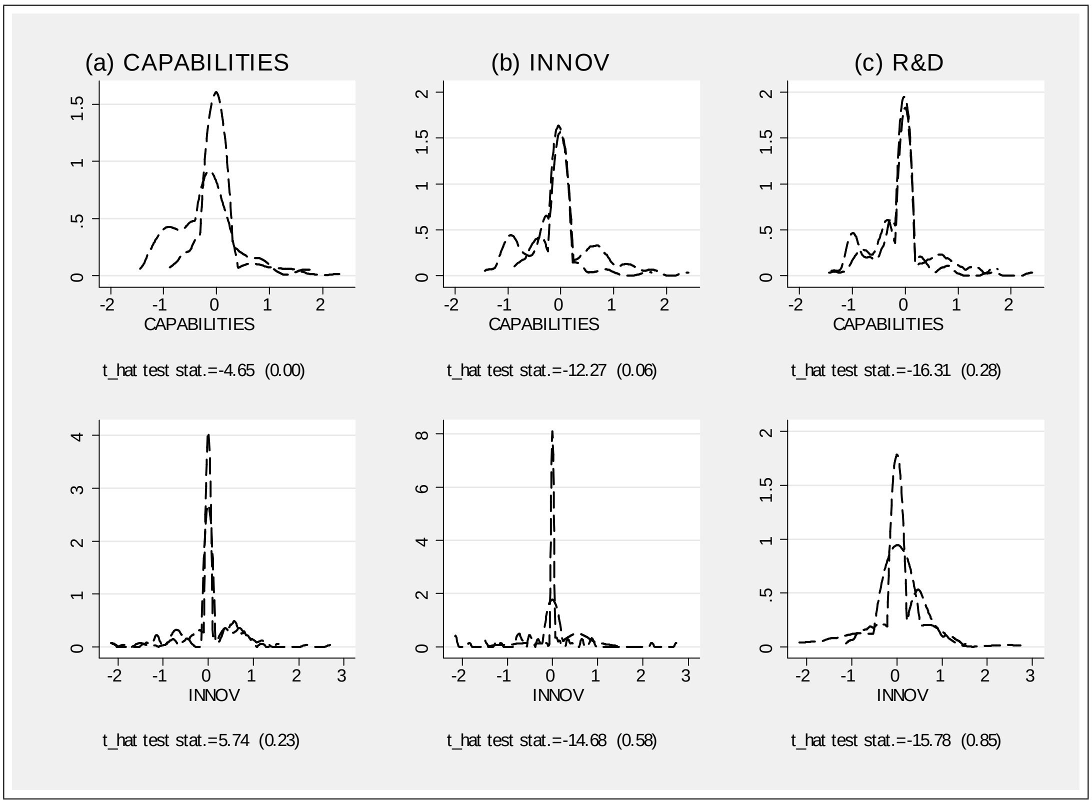

Figure 2: Kernel density estimation of the estimated gradients for selected regional growth determinants. Threshold effects induced by three variables that take values above (solid) and below (dashed) the sample median. Reported values correspond to the Li et al. (2009) test statistic for equality of distributions. Bootstrap p-values in parentheses (399 replications).

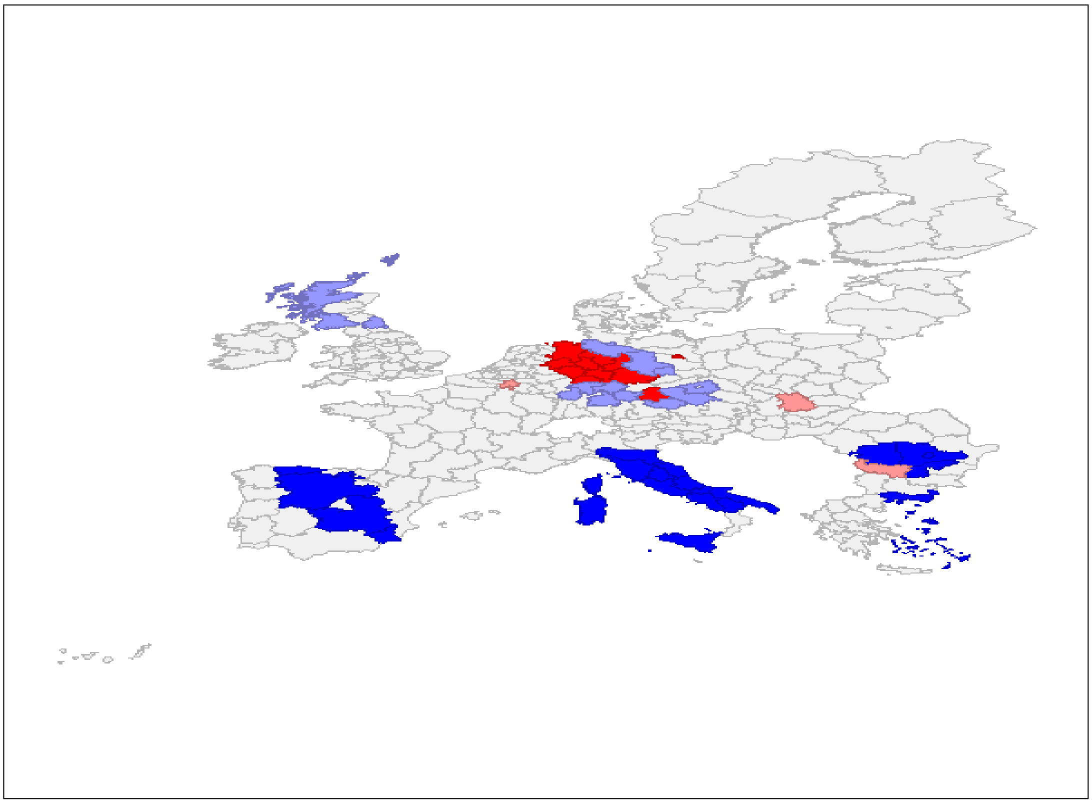

Figure 3: LISA cluster map for the significant partial effects of the share of firms introducing innovations. HH: red; LL: blue; HL: light red; LH: light blue.

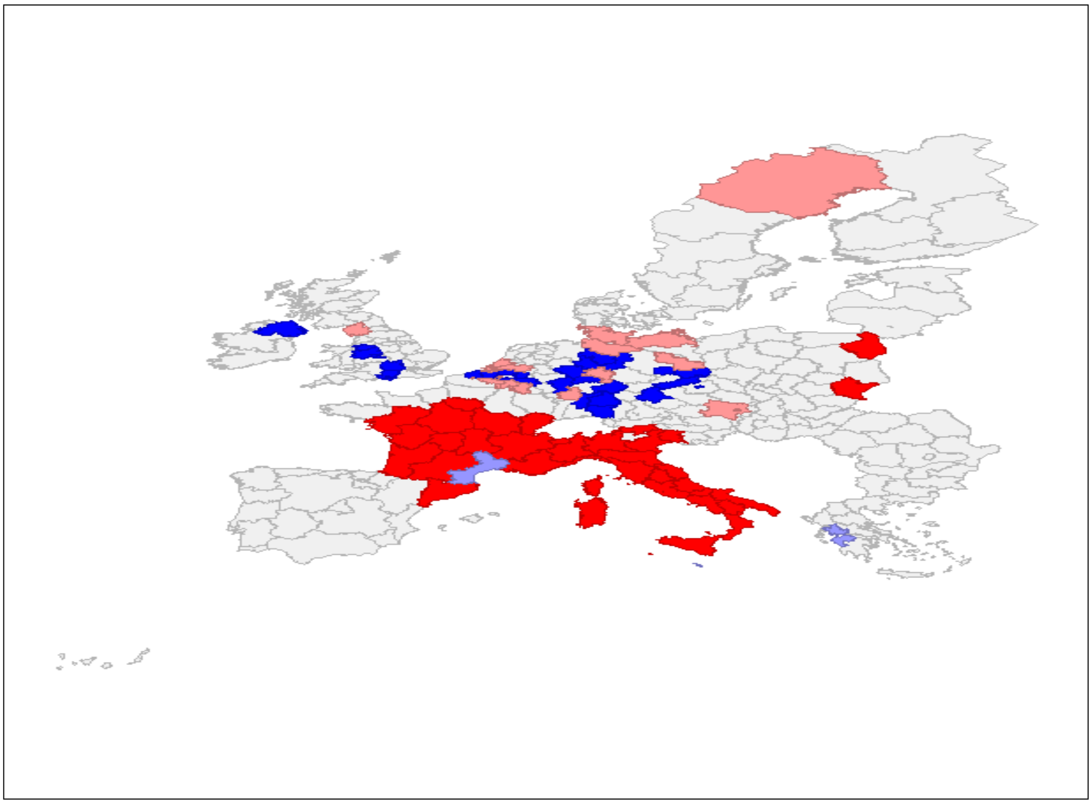

Figure 4: LISA cluster map for the significant partial effects of inward FDI flows. HH: red; LL: blue; HL: light red; LH: light blue.