Laboratoire d'Economie Appliquée de Grenoble

AN 'EXTENDED' KNOWLEDGE PRODUCTION FUNCTION APPROACH TO THE GENESIS OF INNOVATION IN THE EUROPEAN REGIONS

CHARLOT Sylvie ; CRESCENZI Riccardo ; MUSOLESI Antonio

- August 2012 -

JEL CODES : O32 ; R11 ; C14 ; C23

Working Paper GAEL ; 2012-06

Institut National de la Recherche Agronomique - Université Pierre Mendès France
Unité Mixte de Recherche 1215
Domaine Universitaire - BP 47 - 38040 GRENOBLE Cedex 9
Tél. : 33 (0) 4 76 82 54 40 - Fax : 33 (0) 4 76 82 54 55
E-mail : vertier@grenoble.inra.fr - http://www.grenoble.inra.fr

An 'extended' Knowledge Production Function approach to the genesis of innovation in the European regions

Sylvie Charlot[*], Riccardo Crescenzi[†] and Antonio Musolesi[‡]

August 5, 2012

Abstract

The paper looks at the genesis of innovation in the EU regions in order to shed light on the link between innovative inputs (R&D and Human Capital) and the genesis of economically valuable knowledge. The 'traditional' regional Knowledge Production Function (KPF) is innovatively developed in three complementary directions. First, the KPF is 'augmented' in order to control for all possible 'unobservable' and 'immeasurable' time varying factors that influence the genesis of innovation (i.e. localised institutional and relational factors, regional innovation policies). Second, a semi-parametric approach that relaxes any arbitrary assumption on the 'shape' of the KPF is adopted. Finally, the assumption of homogeneity in the impact of R&D and Human Capital is relaxed by explicitly accounting for the differences between 'core' and 'peripheral' regions. The econometric results confirm the importance of accounting for time varying unobserved heterogeneity through the adoption of a "random growth" specification: R&D efforts exert a significant influence on innovation only after controlling for regional specific time varying unobserved factors. In addition the semi-parametric approach uncovers significant threshold effects for both R&D expenditure and Human Capital and highlights a strong complementary between these two factors. However, 'core' regions benefit from a persistent advantage in terms of the 'productivity' of their innovation inputs. This has important implications for the EU innovation policies at the regional level.

JEL classification: O32, R11, C14, C23

Keywords: Innovation, Regions, Knowledge production function, Europe, Semi-parametric models.

[*]GAEL UMR 1215, INRA- University of Grenoble (France). Sylvie.Charlot@grenoble.inra.fr

[†]London School of Economics, Department of Geography & Environment and Spatial Economics Research Centre, London (United Kingdom). r.crescenzi@lse.ac.uk

[‡]CESAER UMR 1041 INRA, Dijon (France). Antonio.Musolesi@dijon.inra.fr

"Because the additive model is not really a very good description of knowledge production, further work on the best way to model the R&D input would be extremely desirable"
Bronwyn H. Hall, Jacques Mairesse, Pierre Mohnen (2010, p. 33)

1 Introduction

Since the seminal contribution by Griliches (1979), the concept of Knowledge Production Function (KPF) has become very popular in the economic literature for the analysis of innovation processes. Two different - and somewhat separate - strands of literature have contributed to the popularity of this approach. On the one hand, micro-level analyses have developed the investigation of the firm-level nexus between innovative input and output and have been recently revamped by the increased availability of Innovation Survey data (Mairesse and Monhen, 2010). On the other hand, the analysis of innovation at the regional level has also made extensive use of 'Regional' KPF in order to assess the contribution of regional inputs to the generation of new local knowledge. This approach extends the 'traditional' firm-level framework *à la* Griliches (1979 and 1986) and Jaffe (1986) in order to account for the role of territorial characteristics and spatial processes (Audretsch and Feldman 1996, Crescenzi et al. 2007 and 2012, O hUallachain and Leslie 2007, Ponds et al 2010), shedding new light on the territorial-level determinants of innovation and their geography.

While both these strands of literature have contributed to a better understanding of different aspects of the process of innovation, more limited attention has been devoted to the synergies between the micro-firm-level and meso-regional perspectives. This paper aims to fill this gap by challenging the micro-level assumptions underlying the specification of regional KPFs. In particular, the paper shows that 'customary' assumptions can be relaxed on a) its specification in terms of 'unobservable' factors; b) its functional form, and c) on the homogeneous (across regions) impact of R&D and human capital on the generation of new knowledge, shedding new light on the process of innovation and its determinants at the territorial level. For this purpose, the paper develops an econometric strategy based on a semi-parametric model with additive unobserved heterogeneity where estimation is based on the Generalized Additive Model framework (GAM, Hastie and Tibshirani, 1990). In this context it is possible to fully control for 'unobserved' and 'unmeasurable' time varying characteristics of the regions (such as localized institutional conditions and relations and regional and local policies) that are overlooked by existing analyses due to the lack of appropriate 'direct' proxies for these factors. In addition, the semi-parametric framework makes it possible to avoid any a-priori assumption on the 'shape' of the KPF (although the Cobb-Douglas specification is customary in the literature, the conceptual section of the paper will show that there is not any theoretical justification for the adoption of this specific functional form). Finally, interaction models (Ruppert et al., 2003, Ch. 12) allow for the identification of heterogeneous relations between innovative inputs and output in different groups of regions in order to test whether these factors show the same 'productivity' in all contexts.

The econometric analysis looks at the regions of the European Union over the 1995-2004 period and confirms the importance to control for time varying heterogeneity: R&D efforts exert a significant influence on innovation only after controlling for other localized factors. Conversely, the impact of Human capital is more robust. In addition the semi-parametric approach uncovers significant threshold effects for both R&D expenditure and Human Capital and highlights a strong complementary between these two factors. Investments in R&D can boost regional innovation only when coupled by a supportive endowment of human capital. In this context, the economically stronger ('core') regions of the EU benefit from a persistent advantage in terms of the 'productivity' of their innovation inputs. In economically disadvantaged 'peripheral' regions, the only predictor for innovative performance remains human capital. However, in these regions, indigenous R&D efforts remain important in order to reinforce their absorptive capacity. These results have relevant implications for EU-level innovation policies and support - in line with other studies adopting a similar approach (e.g. Foddi et al. 2012) - the progressive reduction in the emphasis on R&D investments as a means to promote development in lagging regions in favour of more balanced territorial policies targeting simultaneously human capital and localized institutional factors (Barca, 2009).

The paper is organized as follows. In the next section, the relevant background literature is reviewed in order to develop the foundations for an 'extended' regional knowledge production function approach and identify the most suitable econometric approach for its empirical estimation. Section 3 presents the database and discusses the empirical findings. Section 4 concludes with some policy implications.

2 From the micro-firm level to the regional 'aggregate' specification

Knowledge is not directly observable and it has been traditionally incorporated into standard production functions as an additional separable factor of production proxied by R&D investments (Minisian, 1962; Griliches and Mairesse, 1983; Hall and Mairesse, 1996; Wakelin, 2001;

Ornaghi, 2006). Griliches (1979) proposed a conceptual framework (the Knowledge Production Function approach) suitable to explicitly look at the causal relation between productivity, the unobservable knowledge capital and its observable input (R&D) and output (i.e. patents)[1]. Even if Griliches's knowledge production function was *"at best a very crude reduced-form type relation whose theoretical underpinnings have still to be worked out"* (Griliches, 1990, p.1672), it formed the basis of more sophisticated empirical analyses, estimating simultaneously the R&D, patents and productivity's equations and is now commonly adopted in empirical firm-level studies (Benavente, 2006 ; Crépon et al.,1998; Janz et al. 2004; Jefferson et al. 2006; Lööf and Heshmati, 2002; Lööf and Heshmati, 2006; Van Leeween and Klomp, 2006)[2].

In addition to R&D investments, firm-level studies usually include in their KPFs additional formal (and informal) sources of knowledge and other firm-level characteristics affecting innovative capabilities (Cohen et al.,1989): firm size (Schumpeter, 1912), organizational status, demand pull and technology push (Schmookler, 1962) factors etc. In this context a special role is played by human capital: Very specific skills are necessary to support and enhance the innovative process. In addition the human capital endowment of individual innovative actors is not only directly supportive of their capability to generate new ideas but it also influences their capability to absorb external knowledge (absorptive capacity) in the form of knowledge spillovers.

While the micro-level literature has devoted substantial attention to the analysis of the internal (intra-firm) drivers of innovation, inter-firm knowledge flows have remained relatively unexplored in this strand of literature. Conversely, inter-firm knowledge flows and their diffusion mechanisms have been extensively analyzed in the literature on the innovative performance of 'regions' where firms (and other innovative actors) are 'aggregated' into 'territorial' units of analysis. Firm-level R&D expenditure generates knowledge that can potentially benefit other firms, making public returns to R&D investments greater than private ones.

In a spatial context, the importance of knowledge spillovers impacting upon firms' innovative performance has been linked to agglomeration economies and local institutional factors (including public policies).

Three key sources of agglomeration economies have been identified in the literature: pecuniary externalities linked to the proximity of customers and suppliers; labor market thickness conducive to a better matching between employers and employees and - most relevant for this paper -

[1] After controlling for other relevant factors.

[2] More recently, the use of data obtained from the Community Innovation Survey has allowed the use of a more direct measure of the innovative output, which is a binary variable indicating whether an innovation has been introduced or not (Duguet, 2006; Parisi et al. 2006; Mairesse and Mohnen 2005; Griffith et al. 2006; Mairesse and Robin, 2008, Huiban and Musolesi, 2010, Blanchard et al, 2010).

pure technological externalities. In Marshall's spirit the latter agglomeration externalities result from the ability of people to exchange information and knowledge, depending on their ability to communicate (Charlot and Duranton, 2004).

As far as localized institutional factors are concerned, universities and public research centres influence the 'productivity' of local private R&D in "science-based" sectors by means of localized spin-offs, mobility of employees to and from local universities, and informal exchanges of information through localized social networks (Ponds et al., 2010). In addition, regional innovation policies also influence the innovative performance of firms (Jaffe, 1989; Breandan et al. 2007).

Where both firm-level and regional-level determinants of innovation are jointly taken into account, innovative performance becomes the result of internal firm-level observable (i.e. R&D intensity and Human Capital) and non-observable (e.g. managerial skills, organizational features) characteristics but also of regional-level observable (knowledge spillovers, density) and non-observable (institutional quality, incentives etc.) factors.

2.1 A reduced form equation for the regional Knowledge Production Function

In the light of the factors previously discussed, the firm-level Knowledge Production Function can be 'extended' in order to explicitly account for spatial territorial-level factors:

$$K_{i,t,r} = F(RD_{i,t,r}, HK_{i,t,r}, E_{i,t,r}, V_{i,t,r}) \tag{1}$$

where F is a (real) unknown function, $i = 1, ..., M$ is an index of firms, $t = 1, ..., T$ is an index of time and finally $r = 1, ..., R$ refers to regions; $K_{i,t,r}$, $RD_{i,t,r}$, and $HK_{i,t,r}$ are respectively the innovative output, R&D spending and the level of human capital for the firm i at time t, located in region r. $E_{i,t,r}$ is the external level of knowledge at time t, which firm i benefits from when localized in region r. $V_{i,t,r}$ measures unobservable time varying firm characteristics that facilitate knowledge generation, diffusion and absorption. Some of these characteristics are supposed to be correlated with innovation inputs, such as technological or industrial sector of activity at the firm level. This also includes unobservable time varying characteristics of the region where the firm is localized (such as the mobility of researchers, the effects of social networks etc.) can also be expected to be correlated with innovation inputs.

F is unknown since - as emphasized among others by Hall, Mairesse, and Mohnen (2010) - the innovation process is so complex that there is no a-priori reason for the selection of a linear model. Moreover there is no theoretical justification for the selection of a specific functional form, even though large part of the existing empirical studies is based on a Cobb-Douglas specification.

When looking at the firm-level KPF there is no clear theoretical guidance on how to specify a regional-level 'aggregate' function. The effect of firm-level R&D spending and human capital on innovative output depends on both the unobservable (time-varying or idiosyncratic) firm characteristics and on the 'external' regional environment.

Following this line of reasoning, regional innovative output $K_{r,t}$ which is defined as $K_{r,t} = \sum_{i=1}^{M} K_{i,t,r}$, can be assumed to be a (unknown) function, G,of the regional level of human capital and R&D spending that also depends on non-observable regional characteristics:

$$K_{r,t} = G(RD_{r,t}, HK_{r,t}, U_{r,t}) \tag{2}$$

where $RD_{r,t}$, and $HK_{r,t}$ are R&D spending and human capital at the regional level. They are defined as: $RD_{r,t} = \sum_{i=1}^{M} RD_{i,t}$ and $HC_{r,t} = \sum_{i=1}^{M} HK_{i,t}$. As previously, $U_{r,t}$ can be considered as a proxy for all unobservable (or unmeasurable) factors influencing regions' innovative performance and again some of these characteristics may also be correlated with innovation inputs. This correlation can be the result of the non random selection of the location of both R&D investments and human capital. In the same way as firms accurately select the location of their R&D investments, high skilled workers choose non randomly where to look for jobs. As a consequence, whereas patenting depends on R&D spending, human capital and non-observable (and possibly time varying) characteristics, - at the same time - R&D spending and human capital also depend on the latter set of characteristics (e.g. agglomeration, regional industrial structure, infrastructure, etc.) that determine their location decisions.

When firm-level variables are aggregated into regional-level indicators there is not any a-priori theoretical reason as for why the link between innovative input and output should keep the same functional form. The knowledge production function can, in principle, exhibit different shapes at the firm and at the regional 'aggregate' level. The relationship observed at the firm level can change at the regional level when taking into account the regional conditions for knowledge circulation. If local firms are very differentiated in terms of their capability to absorb local knowledge the functional form can change substantially when aggregating firm-level observations. As a consequence, the level of aggregation could impact upon the returns to scale of R&D and human capital on innovation production.

In addition, regional-level innovation also depends on inter-regional spillovers: innovative activities pursued in neighboring regions influence local innovative performance. The impact of inter-regional spillovers depends on local absorptive capabilities but also on the amount and nature of external innovation in terms of technological and cognitive congruence (Boschma 2004;

Bode 2004).

As a consequence, when estimating a KPF at the regional level it is necessary to take into account the conditions in neighboring regions, in terms of both R&D and human capital. It is therefore necessary to introduce the R&D spending and human capital in neighboring regions as an argument of the regional knowledge production function. As a consequence, the regional knowledge production function to be estimated takes the following general form:

$$K_{r,t} = K(RD_{r,t}, HK_{r,t}, WRD_{r,t}, WHK_{r,t}, U_{r,t})$$

where $WRD_{r,t}$ and $WHK_{r,t}$ are respectively R&D and human capital level in the regional neighborhood. W is a generic inverse-distance matrix (allowing the impact of the activities pursued in neighboring regions to decay with distance). This regional knowledge production function can be written more compactly as:

$$K_{r,t} = K(X_{r,t}, U_{r,t}) \tag{3}$$

where $X_{r,t}$ is a vector of explanatory variables: $X_{r,t} = (RD_{r,t}, HK_{r,t}, WRD_{r,t}, WHK_{r,t})$.

2.2 The econometric framework

In order to model the regional KPF, the starting point is the following parametric additive linear model:

$$K_{r,t} = X_{r,t}\beta_r + U_{r,t} \tag{4}$$

where $K_{r,t}$ is innovative output, $X_{r,t}$ is a $1 \times K$ line vector of time variant observed regional variables (innovation inputs), β_r is the $K \times 1$ vector of individual slopes and $U_{r,t}$ is an error term. $r \in \Gamma$, and Γ is the set of cross-section units $\Gamma = \{1, 2, ..., N\}$ and $t \in \Lambda$=$\{1, 2, ..., T\}$ indicates time series observations. We also assume that the unobservable term $U_{r,t}$ can be decomposed as:

$$U_{r,t} = Z_t\alpha_r + u_{r,t} \tag{5}$$

where Z_t is a $1 \times J$ vector of nonrandom aggregate time variables which can be freely correlated with the explanatory variables $X_{r,t}$ and α_r is the associated $J \times 1$ vector of individual coefficients; $u_{r,t}$ is an error term supposed to be orthogonal to the $X_{r,t}$. The KPF can be written:

$$K_{r,t} = Z_t\alpha_r + X_{r,t}\beta_r + u_{r,t} \tag{6}$$

For such a model, the strict exogeneity assumption on the explanatory variables (see Wooldridge, 2005, 2010) is

$$E\left(u_{r,t} \mid X_{r1}, ..., X_{rT}, \alpha_r, \beta_r\right) = 0, \qquad\qquad t = 1, ..., T$$

which follows from the conditional mean assumption

$$\begin{aligned} E\left(K_{r,t} \mid X_{r1}, ..., X_{rT}, \alpha_r, \beta_r\right) &= E\left(K_{r,t} \mid X_{r,t}, \alpha_r, \beta_r\right) \\ &= Z_t \alpha_r + X_{r,t} \beta_r, \qquad\qquad t = 1, ..., T \end{aligned}$$

where the first equality gives the strict exogeneity interpretation: conditional on $(X_{r,t}, \alpha_r, \beta_r)$, the covariates from the other periods do not affect the expected value of $K_{r,t}$. However this does not restrict the correlation between (α_r, β_r) and $(X_{r1}, ..., X_{rT})$, thus taking into account the endogenous nature of innovation inputs (i.e. R&D and human capital).

In this spatial context, the concept of *selection on unobservables* (Heckman and Hotz, 1989) is particularly relevant: For example, in the (time invariant) unobserved effects model (with $Z_t = 1$ and $\beta_r = b$), the unobserved effect, α_r, can capture both first nature (or natural) advantages (i.e. physical geography) and second nature advantages (i.e. 'the geography of distance between economic agents'; Overman et al., 2001, p. 2). If aggregate time effects are included into the model (two way fixed effects with $Z_t = (1, \lambda_t)$, $\alpha_r = (\alpha_{r1}, \alpha_2)$ and $\beta_r = b$), then our specification also accounts for the possible dependence between innovation inputs and unobserved time related factors, which are assumed to affect homogeneously the dependent variable. Finally, a more general specification considers individual time trends, in addition to the individual and time fixed effects, as in Papke (1994) or Friedberg (1998): $Z_t = (1, \lambda_t, t)$, $\alpha_r = (\alpha_{r1}, \alpha_2, \alpha_{r3})$ and $\beta_r = b$. These region-specific time trends also account for region-specific time-related unobservable factors: Firms may locate their R&D activities where they can find all factors supportive to innovation (i.e. other innovative firms, universities, localised institutional factors etc.).

Furthermore, when analysing the innovative performance of regions, it is plausible to expect significant differences across regions in terms of returns of R&D and human capital, so that the slope coefficient, β_r, should be allowed to vary across regions.[3] In addition, the possible correlation between β_r and $X_{r,t}$ makes model (6) a *correlated random-coefficient model.* Heckman and Vytlacil (1998, p.975), referring to the return of schooling on wages, state *"It is plausible that*

[3] This has motivated interest in random-coefficient models (Hsiao, 2003, chapter 6) where the slope coefficients can vary with the cross sectional units. However, at the same time, the underlying assumption that such coefficients are independent of the regressors may be too much restrictive.

the return to schooling declines with the level of schooling". A similar reasoning can be applied in the KPF context with respect to the return of both human capital and R&D on the creation of new knowledge. For example a critical mass of R&D and/or human capital might be necessary to produce a significant impact on innovation (as in the Schumpeterian perspective), making the region-specific effects of R&D and/or human capital on innovative output potentially dependant upon the level of these inputs.

Fixed effects estimators of the kind presented before are consistent not only when unobserved heterogeneity is correlated with the explanatory variables but also when considering a correlated random-coefficient model allowing individual slopes potentially correlated with the explanatory variables (Hsiao, 2011; Wooldridge 2005). Indeed, under the random coefficient assumption we can write $\beta_r = b + d_r$ where d_r is a random vector representing the unit-specific deviation from the average, b, which has zero mean by definition. Wooldridge (2005) shows that, under the random coefficient assumption, a sufficient condition for consistency of fixed-effects type estimators, is that:

$$E\left(\beta_r \mid \ddot{X}_{r,t}\right) = E\left(\beta_r\right) = b \tag{7}$$

where $\ddot{X}_{r,t}$ is the vector of regressors after a "general fixed effects transformation". It is the first equality that gives the condition for consistency: it says that β_r is mean independent of $\ddot{X}_{r,t}$. It is worth noticing that such a condition allows the slopes β_r to be correlated with $X_{r,t}$ via the the permanent components. A quite natural assumption is that $X_{r,t}$ has the same deterministic component as for $K_{r,t}$ and, consequently, if we assume that the DGP (Data Generating Process) for $K_{r,t}$ is characterized by individual and temporal effects and individual trends $Z_t = (1, \lambda_t, t)$, $\alpha_r = (\alpha_{r1}, \alpha_{r2}, \alpha_{r3})$, it seems natural to write $X_{r,t} = f_r + m_t + g_r t + r_{r,t}$ where $r_{r,t}$ is an idiosyncratic component and $f_r + m_t + g_r t$ is the permanent component. The condition for consistency $E\left(\beta_r \mid \ddot{X}_{r,t}\right) = E\left(\beta_r\right)$ holds if

$$E\left(\beta_r \mid r_{i1}, ..., r_{iT}\right) = E\left(\beta_r\right) \tag{8}$$

which means that β_r is mean independent of the idiosyncratic component of $X_{r,t}$ so that β_r can be arbitrarily correlated with the permanent component (f_r, m_t, g_r). This is very important in our empirical framework because it allows for the region-specific effect of innovation inputs (R&D and Human capital) on innovative performance to be correlated with long-term innovative conditions (f_r), the region-specific growth rates in such conditions (g_r) and the common deviations in relation to factors such as oil prices modification, global recession etc. (m_t) that can affect all regions simultaneously.

In the next section, the parametric analysis is extended by considering a semi-parametric model with additive unobserved heterogeneity that *explicitly* allows the effect of innovation inputs to vary freely with their level. In its simplest form, i.e. setting $X_{r,t} = (RD_{r,t}, HK_{r,t})$, such a model can be written as

$$K_{r,t} = Z_t\alpha_r + f_1(RD_{r,t}) + f_2(HK_{r,t}) + u_{r,t} \tag{9}$$

where f_1 and f_2 are smooth functions. The estimation and identification of equation (9) can be obtained by considering the Generalized Additive Model framework methodology (GAM, Hastie and Tibshirani, 1990).[4] The estimation is carried out by maximizing a penalized likelihood by penalized iteratively reweighted least squares (P-IRLS) (Wood, 2004) [5] while the choice of the basis to represent the smooth terms and the selection of the smoothing parameters follows Wood (2003, 2006ab, 2008). [6] More general specifications can be obtained by considering a continuous-by-continuous interaction which does not impose additivity or a factor-by-continuous interaction (Ruppert et al., 2003), making it possible to introduce a greater level of cross sectional heterogeneity concerning the effect of RD and HK.

3 Econometric analysis

3.1 Data

The analysis covers the entire EU15 for the period 1995-2004. Data on innovative output (patents) and explanatory variables on its determinants are available from Eurostat. The analysis is based on a mix of NUTS1 and NUTS2 regions, selected in order to maximize their homogeneity in terms of the relevant governance structure under the constraint of data availability. The unit of analysis with the greatest relevance in terms of the institutions supporting the process of innovation and

[4] A priori, the fixed effects parameters α_r can be treated as nuisance terms to be eliminated with a transformation or as parameters to be estimated. Considering the α_r as nuisance terms allows to reduce significantly the number of parameters to be estimated but, at the same time, can produce serious identification problems (Su and Ullah, 2010). Since the relatively high number of time series observations (T=10) of our data set allows to recover all the parameters (and functions) of interest when the level equation is estimated (and thus the fixed effects parameters have to be estimated), we follow this approach, as in Mammen et al. (2009) or Ordás Criado et al. (2009). This also allows the identification of all the components of (9).

[5] It is worth noticing that the additive structure of this specification allows us not to rely on a local linear approximation of the model (Su and Ullah, 2006, 2010).

[6] The smoothing parameters are selected directly using the so-called 'outer iteration' (Wood, 2008) which has been shown to be computationally efficient and stable.

its diffusion and which may be taken as a target area for innovation policies by the national government and/or the European Commission – was selected for each country. Consequently, the analysis is based on NUTS1 regions for Belgium, Germany[7] and the United Kingdom and NUTS2 for all other countries (Austria, Finland, France, Italy, the Netherlands, Portugal, Spain, Sweden). Countries without equivalent sub-national regions (Denmark[8], Ireland, Luxembourg) are excluded a priori from the analysis.[9]

We measure the regional innovation level by the regional number of patents per capita per year. The analysis is based on homogenous comparable data for regions belonging to all EU countries and relies on European Patent Office data as recorded by Eurostat in its regional database. Regional R&D expenditure is provided by Eurostat and corresponds to the share of regional GDP spent in R&D by public and private institutions, whatever the sector. Finally human capital is proxied by the regional share of workers with tertiary education or higher (ISCED76 classification levels 5-7).

In order to assess the spillovers effects from R&D and human capital in neighboring regions, spatially lagged variables are computed by means of inverse of Euclidian distance matrices (see Appendix 2 for details) as customary in the literature on spatially-mediated knowledge spillovers (Moreno et al. 2005).

3.2 The role of 'unobserved' and geographical (time varying) factors

The previous section has shown how the estimation of a fully specified Regional Knowledge Production Function requires to account for a number of 'unobservable' factors that influence the generation of new knowledge . As a consequence, the 'basic' specification of our empirical model explicitly includes these factors while adopting the log-log specification (called "Cobb-Douglas" to

[7] The NUTS2 level corresponds to *Provinces* in Belgium and to the German *Regierungsbezirke*. In both cases these statistical units of analysis have little administrative and institutional meaning. For these two countries the relevant institutional units are *Régions* and *Länder*, respectively, codified as NUTS1 regions. The lack of correspondence between NUTS2 level and actual administrative units accounts for the scarcity of statistical information on many variables (including R&D expenditure) below the NUTS1 level for both countries.

[8] Even if Denmark introduced regions above the local authority level on January 1^{st} 2007 in line with the NUTS2 classification, regional statistics are not available from Eurostat.

[9] As far as specific regions are concerned, no data are available for the French Départments d'Outre-Mer (FR9). Trentino-Alto Adige (IT31) has no correspondent in the NUTS2003 classification. Due to the nature of the analysis, the islands (PT2 Açores, PT3 Madeira, FR9 Départements d'Outre-mer, ES7 Canarias) and Ceuta y Melilla (ES 63) were not considered due to the problems with the computation of the spatially lagged variables.

recall the 'traditional' production function) as customary in the existing regional-level studies.[10] In this preliminary specification the set of explanatory variables is constrained to include only R&D and human capital, i.e. the direct input of new knowledge generation

$$\log(K_{r,t}) = Z_t\alpha_r + \beta_1 \log(RD_{r,t}) + \beta_2 \log(HK_{r,t}) + u_{r,t}$$

and the results based on alternative definitions for the deterministic component $Z_t\alpha_r$ are compared:

- $Z_t = 1$, 'one way fixed effects'
- $Z_t = (1, \lambda_t)$, $\alpha_r = (\alpha_{r1}, \alpha_2)$, 'two way fixed effects'
- $Z_t = (1, \lambda_t, t)$, $\alpha_r = (\alpha_{r1}, \alpha_2, \alpha_{r3})$, 'random growth'

In the (time invariant) unobserved effects model ($Z_t = 1$), the unobserved effect α_r captures "the fundamental determinants of the spatial variation of per capita income (...): First nature geography; the second nature geography of access to markets, suppliers, and ideas: and third, the effects of social infrastructure" (Overman et al.2003, p.11). In order to take into account the evolution over time of all these unobserved common factors, two additional sets of controls are introduced. First, a set of year dummies included: under the restrictive assumption that the unobserved common time factors affect homogeneously (i.e. in the same fashion in all regions) the dependent variable it is possible to consistently estimate the slope parameters by adopting a two-way fixed effects model. Finally, a more general specification considers individual time trends, in addition to the individual and time fixed effects, as in Papke (1994) or Friedberg (1998): these trends account for the dynamic evolution of unobservable factors influencing the innovative capabilities of each individual region.

[10] It should be noted that, in firm-level empirical analyses, the specification of the KPF generally depends on the nature of the available proxy for innovative output. For example, Crépon et al. (1998) specify an ordered probit for the share of innovative sales (or a poisson regression for patents' counts), while Griffith et al. (2006) and Blanchard et al. (2010) use CIS data and propose a probit for the binary variable indicating whether an innovation has been introduced or not. In both cases, the latent (not observed) innovative output variable depends on the latent R&D effort in logs and on a vector of other firms characheristics affecting their capacity to innovate. Other formulations have also been proposed. Duguet (2006) and Huiban and Musolesi (2010) use extensively CIS data and replace, as explanatory variable, the continuos latent R&D effort with its binary couterpart indicating whether a firm has pursued R&D activities in the recent past.

The estimation results are shown in Table 1 where robust standard errors are reported[11]. Columns (1), (2) and (3) report “one way”, “two way” and “random growth” coefficients respectively. Since the three specifications discussed above are nested we use sequentially the F test to choose among them. The F-Test strongly supports the ’random growth’ specification.

[Table 1 about here]

The results show that significant discrepancies in the estimated coefficients emerge for different specifications of the unobservable part of the model, both in terms of magnitude and significance of the estimated coefficients. When controlling only for time-invariant regional characteristics (column 1) the generation of innovation is dominated by human capital endowment while regional R&D investment doesn’t seem to play a significant role. Regional fixed effects are able to control for long-term structural conditions of the local economy allowing human capital to emerge as a fundamental driver for innovative performance. However, after controlling for the dynamics of the ’unobservable’ factors over time as in the ’random growth’ specification, both R&D and Human Capital are significant and show elasticities of a similar magnitude (0.60 and 0.51 respectively). The ’random growth’ specification is most suitable in order to fully control not only for the role of ’first’ and ’second’ nature geography that are relatively stable over time but also for the effects of the ’social infrastructure’ supportive of the process of regional innovation. A successful process of innovation depends on “localised structural and institutional factors that shape the innovative capacity of specific geographical contexts” (Iammarino 2005, p.499) as highlighted by the systems of innovation approach (Lundvall 2001), regional systems of innovation (Cooke 1998) and learning regions (Morgan 2004). Building on such contributions we support the idea that the variety of factors brought into “focus” by the systemic approach to the genesis of innovation, while ‘immeasurable’ and resistant to any generalisation when used as linear predictors for innovative performance, can be ‘fully controlled for’ in the ‘random growth’ specification of the Regional KPF. Only after controlling for their intrinsically ‘dynamic’ influence on regional innovation, the role of R&D can emerge as statistically significant. This is in line with the evidence produced by Rodriguez-Pose and Crescenzi (2008) when looking at the impact of R&D investment on regional growth in the EU: R&D expenditure can produce ”very different results, according to the possibility of every region of benefiting from (. . .) favorable underlying socioeconomic conditions” (p.63).

[11] Driscoll and Kraay’s (1998) standard errors which are robust to heteroskedasticity, serial and spatial correlation.

With the 'random growth' specification, the elasticity of patents with respect to R&D is estimated at 0.6 which is coherent with firm level analyses. Indeed, Griliches (1990), summarizing previous results reports that the elasticity of patents with respect to R&D is between 0.3 and 0.6, while Blundell et al. (2002) report a preferred estimate of 0.5 (see also Klette and Kurtom, 2004). Regional-level studies based on KPFs produced comparable results: OhUallachain and Leslie (2007) report an elasticity of patents to R&D of 0.575 for US states, Ponds et al. (2010) calculated an elasticity of 0.429 to university R&D and 0.373 to private R&D for regions in the Netherlands and Foddi et al. (2012) find an elasticity of 0.42 when looking at the entire EU-27 in 'regional' KPF framework. Similar results are reported in the literature as far as the elasticity with respect to Human Capital is concerned: Foddi et al. (2012) report an elasticity of patenting to regional human capital endowment of 0.65 (in line with the results presented in this paper).

3.3 From the Cobb-Douglas to a non-constrained relation

The lack of a clear theoretical guidance to select and 'justify' the most appropriate functional form is an important constraint for any analysis of regional innovative performance. As a consequence, we relax any assumption on the shape of the regional KPF and estimate the model without imposing a specific parametric specification as detailed in section 2.2. Once again, the results of the tests clearly favour the random growth specification[12], i.e. $Z_t = (1, \lambda_t, t)$, $\alpha_r = (\alpha_{r1}, \alpha_2, \alpha_{r3})$ .In what follows we provide the plot of the estimated functions with their standard errors (when a variable has not a significant effect, it will be highlighted in the text) while detailed "parametric" results on the EDF (Estimated Degrees of Freedom) and p values are available in Appendix 1, for all specifications

The effect of human capital and R&D The impact of regional human capital and R&D expenditure on regional innovation is estimated first :

$$K_{rt} = Z_t\alpha_r + f_1(RD_{r,t}) + f_2(HK_{r,t}) + u_{r,t}$$

Figure 1 represents the estimated functions with their standard errors.

[12] We follow Wood (2006a) and use an approximate F test.

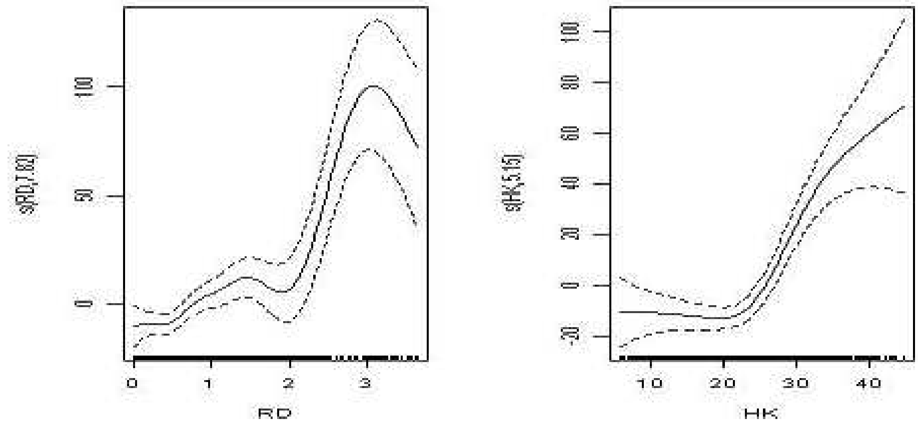


FIGURE 1

These results confirm that, although parametric estimates are robust, they hide an important part of the story on the relationship between innovation and its key inputs (R&D and Human Capital): their link is non-linear and shows a shape that deserves particular attention. The left-hand side section of Figure 1 shows the impact of R&D expenditure on patenting with the black line visualising the value of the estimated coefficient and dotted line the corresponding confidence interval. For low levels of regional R&D Expenditure the impact on patenting is positive although relatively limited. However, above a first threshold (i.e. above 1.8%) the impact of additional investments is much less significant and potentially detrimental: it is only after the second threshold (above approximately 2%) that R&D expenditure maximizes its impact on innovation. This positive and significant relationship holds true until an additional threshold is reached (just below 3%). For very high levels of R&D expenditure the confidence interval becomes too broad to suggest a significant relationship with patenting. Other factors interact with R&D expenditure in explaining the success of the top innovative EU regions where Systems of Innovation conditions have been shown to be of paramount importance for the productivity of R&D activities (Crescenzi et al. 2007 for the EU and the US and 2012 for emerging countries; Iammarino 2005; Lundvall 2001).

For human capital (right-hand side part of figure 1), the slope highlights a single fundamental threshold at approximately 20%. Below this threshold there is no significant impact of human capital on innovation, above this level the impact becomes highly positive and significant (narrow confidence interval). It is only for very high levels of human capital intensity (above approximately 35%) that the strength of this relationship becomes less significant and the confidence interval broadens up again, in line with the previous results for R&D for the most innovative regions in the EU.

However, innovation theory suggests that R&D investments and human capital are strongly complementary in their contribution to innovation. As a consequence, Figure 2 shows the plot for the estimation of the continuous by continuous interaction (that does not impose additivity as in Ruppert et al., 2003) between R&D and Human Capital:

$$K_{rt} = Z_t\alpha_r + f(RD_{r,t}, HK_{r,t}) + u_{r,t}$$

The estimation generates the results visualised in Fig.2

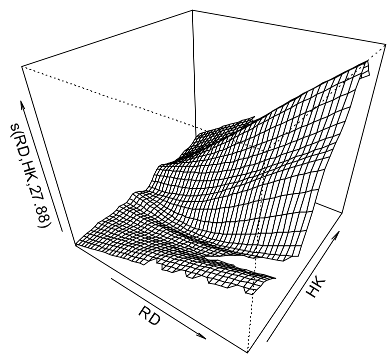


FIGURE 2

The surface in the 3-D plot (Figure 2) shows how innovative output (Z-axis) responds to simultaneous changes in R&D (Y-Axis) and Human Capital (X-axis). For low levels of regional human capital endowment the 'productivity' of R&D investments (of whatever size) remains limited. This confirms the risk of the 'cathedrals in the desert' scenario (Crescenzi and Rodriguez-Pose 2009; Midelfart-Knarvik and Overman 2002) whereby R&D investments are concentrated – for example because of policy incentives in favour of lagging areas – in regions lacking an appropriate receptive environment in terms of human capital. The local mismatch between R&D and skilled labour persistently hampers innovation. Conversely, the plot shows that for higher levels of human capital intensity (after an initial threshold) the impact of R&D investments on innovation is positive and sharply increasing . R&D investments can boost innovation only where appropriate complementary skills are available locally to support knowledge generation and absorption.

Regional spillovers The origin of knowledge flows can be local, but they can also be generated outside the borders of the region under analysis, as “there is no reason that knowledge should stop spilling over just because of borders, such as a city limit, state line or national boundary” (Audretsch and Feldman, 2004, p.6). Knowledge produced in one region spills over into another, influencing its innovative performance (Moreno et al. 2005) and proximity facilitates faster diffusion (Sonn and Storper 2008). Inter-regional spillovers are proxied by a specific variable i.e. the distance-weighted R&D investments pursued in neighboring regions (detail on the computation of the variable in Appendix 2) following Crescenzi et al. 2007. R&D investments pursued in neighboring regions impact upon local innovative performance by means of spatially-mediated diffusion processes that 'augment' locally available knowledge input.

Adding regional spillover effects and estimating the equation:

$$K_{rt} = Z_t\alpha_r + f_1(RD_{r,t}) + f_2(HK_{r,t}) + f_3(WRD_{r,t}) + f_4(WHK_{r,t}) + u_{r,t}$$

provides us with results that are graphically summarized in Figure 3:

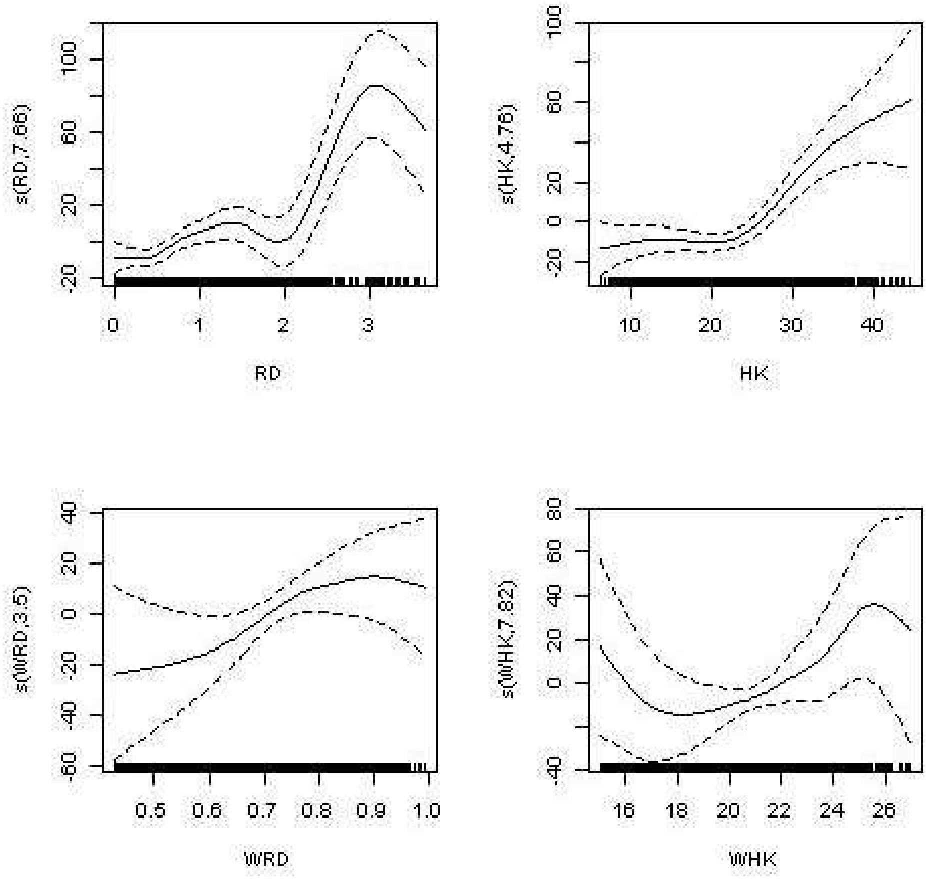
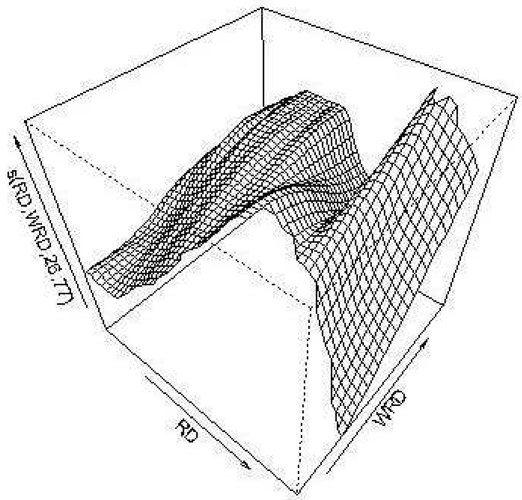


FIGURE 3

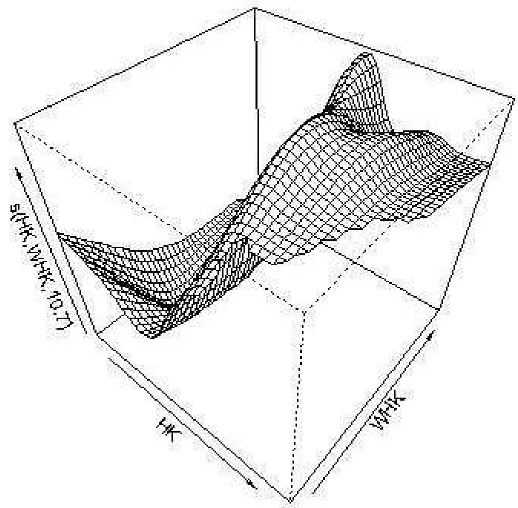
The upper part of figure 3 visualises the estimation results for the impact of R&D and Human Capital on innovation. These results are equivalent to those shown in figure 1. The lower part of the figure shows, instead, the estimations for spillovers generated by R&D activities (WR&D - left-hand side) and Human Capital (WHK - right-hand side) in neighboring regions. Spillovers from R&D activities pursued in neighboring regions have a positive impact on local innovative performance. The spatially-lagged R&D expenditure has a positive and significant impact on innovative output. However, for this effect to be positive and significant (confidence interval reasonably narrow) the exposure to spillovers needs to be sufficiently high. However, above a certain threshold the confidence interval broadens up again and the curve becomes concave. These results suggest – in line with the existing literature (Bottazzi and Peri 2003; Crescenzi et al. 2007; Greunz 2003; Moreno et al. 2005) – that inter-regional spillovers are an important component of the territorial dynamics of innovation in the EU. However, their impact on local dynamics depends (as will be discussed below with interaction terms) on internal absorptive capabilities. In addition, the limited impact for lower levels of exposure suggests that peripheral regions might not be able to rely on inter-regional spillovers in order to reinforce their innovative performance. In this sense peripherality becomes a source of structural disadvantage for innovative performance. The results for human capital (although generally less statistically significant) seem to point in a similar direction: regions surrounded by a neighborhood poorly endowed in terms of human capital (low level of WHK) tend to produce less innovation (negative impact). For higher concentrations of human capital in neighboring regions the effect on internal innovation becomes positive and significant. However, this effect tends to vain for proximity to the big centres of human capital accumulation (very high value of WHK): these hotspots tend to generate a shadow effect in their neighbours attracting their human capital (for high level of WHK the impact is non-significant and tends to become negative).

Finally, the non-additive specification presented in the previous section is re-estimated by interacting local R&D and HK respectively with R&D and Human Capital in neighboring regions. This set of interactions is particularly relevant on both analytical and policy grounds. As discussed above, there is now consensus in the literature on the relevance of spatially-bound inter-regional spillovers for the genesis of innovation in the EU. However, the analysis of the indigenous factors conditioning their impacts remains to be further explored. These interaction terms make it possible to test the indirect effect of both R&D and Human Capital investment as components of the capability of the regional economy not only to produce (direct impact) but also to absorb (indirect impact) external knowledge flows to be 'translated' into innovation. The assessment of the contribution of R&D and Human Capital to local absorptive capabilities is highly relevant to innovation policies in the less innovative areas of the Union: even where a 'critical mass' is not be in place, reinforcing both R&D and Human Capital investments might still be justified as a means to support the absorption of external knowledge. Cohrently with these considerations the following equation is estimated:

$$K_{rt} = Z_t\alpha_r + f_1(RD_{r,t}, HK_{r,t}) + f_2(RD_{r,t}, WRD_{r,t}) + f_3(HK_{r,t}, WHK_{r,t}) + u_{r,t}$$



4.A



4.B

FIGURES 4

The first graph (left-hand side) visualises the results for spillovers from R&D activities and suggests that the impact of localised knowledge spillovers is highly differentiated depending on the local level of R&D spending.[13] For low levels of internal R&D expenditure external R&D activities seems to be able to play a compensatory role and exert a positive influence on local innovation. However, for intermediate levels of internal spending a shadow effect seems to prevail. The impact of a highly innovative neighborhood turns negative where internal expenditure remains too low to counterbalance centripetal forces that might suck innovation away from the local economy in favour of its neighborhood. Finally, when internal R&D is large enough, the relationship between regional patenting and the interaction term becomes again positive: highly innovative regions surrounded by other technologically dynamic regions maximize innovative output due to the strong complementarities between internal and external knowledge flows.

For Human Capital (right-hand side graph) the story is more straightforward. For low levels of internal human capital the shadow effect tends to prevail: if a critical mass is not in place skilled individuals tend to migrate in neighboring regions (Iammarino and Marinelli, 2010). Conversely, where the internal endowment is high enough, the interaction between internal and external human capital becomes a positive driver for innovation. However, this positive impact tends to vain where human capital concentration is very high: for top human capital regions the

[13] The estimated interaction term f_1 is not reported here since it is very similar to that estimated in the previous section.

geographical proximity of other highly endowed centres might create a 'competition for skills' that turns out to be detrimental for innovation (Salt, 2002).

3.4 A 'known' source of heterogeneity: core vs. peripheral regions

A final step in the analysis of the knowledge generation process in the EU regions, is to test whether the territorial dynamics unveiled above work in the same way in all regions or if specific sub-groups or categories of regions show different innovative pattern. In other words it is possible that the EU regional Knowledge Production Function shows heterogeneous effects of R&D (and HK) on Innovation. By testing this potential heterogeneity on the basis of the à priori knowledge on its source we can estimate a more realistic KPF model.

Following this line of reasoning the final section of the analysis aims to test whether regions eligible for Objective 1 support by the EU structural funds - i.e. the most disadvantaged regions in Europe (or 'Convergence Regions' in the current programming period, with a GDP per capita below 75% of the EU average) – show 'heterogeneous' dynamics with respect to those identified for the whole Europe. In order to empirically test this hypothesis we estimate a model where the continuous variables R&D and HK interact with a dummy variable taking the value 1 if a region is eligible for Objective 1 support and 0 otherwise (binary by continuous interaction, Ruppert et al., 2003, chapter 12) thus obtaining two distinct nonparametric functions for each explanatory variable. Letting:

$$O1_r = \begin{cases} 1 \text{ if region } \in \text{ objective 1 group} \\ 0 \text{ if region } \notin \text{ objective 1 group} \end{cases}$$

and estimate:

$$K_{rt} = Z_t\alpha_r + f_{1O1_r}(RD_{r,t}) + f_{2O1_r}(HK_{r,t}) + f_{3O1_r}(WRD_{r,t}) + f_{4O1_r}(WHK_{r,t}) + u_{r,t}$$

Figure 5 shows the estimation results for non-Objective 1 regions (Left-hand side) and for regions eligible for Objective 1 support (Right-hand side):

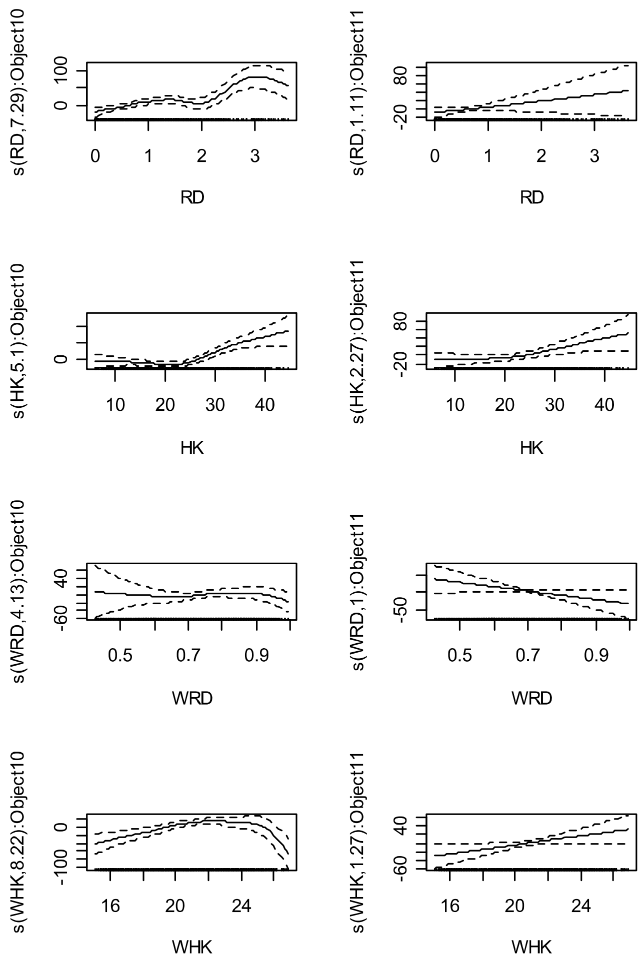


FIGURE 5

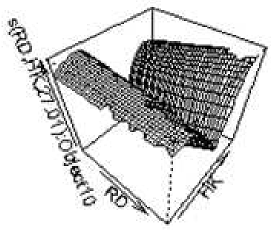
Figure 5 confirms the hypothesis of heterogeneous effects in 'core' and 'peripheral' areas of the Union. The estimated results show that the evidence discussed so far is reinforced when restricting the sample to the most dynamic regions of the EU: results for the impact of R&D (first graph on the left of figure 5), Human Capital (second graph on the left) and Spillovers from R&D and HK (third and fourth graphs on the left) remain unchanged. On the contrary, in Objective 1 regions the only driver that remains positive and statistically significant (narrow confidence interval) is Human Capital (HK – Second graph, right-hand side part of Figure 5). Human capital is the key driver for innovation in lagging regions and, interestingly enough, for this group the threshold effect is less accentuated than in the whole sample. Investing in

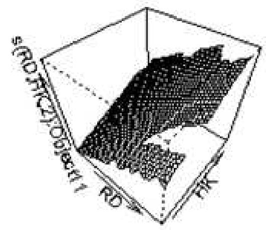
Human Capital in Objective 1 regions is beneficial to innovation irrespective of the their initial endowment and the effect is relatively linear.

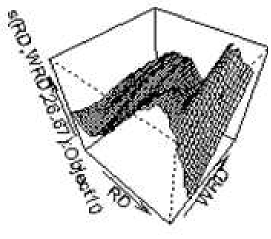
Finally, a more general model is estimated by interacting the binary variable $O1_i$ with the two dimensional splines previously introduced, that is:

$$K_{rt} = Z_t\alpha_r + f_{1O1_r}(RD_{r,t}, HK_{r,t}) + f_{2O1_r}(RD_{r,t}, WRD_{r,t}) + f_{3O1_r}(HK_{r,t}, WHK_{r,t}) + u_{r,t}$$

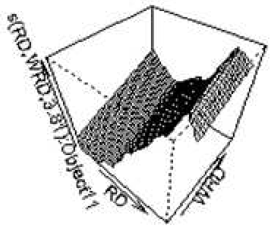
and obtaining Figure 6:

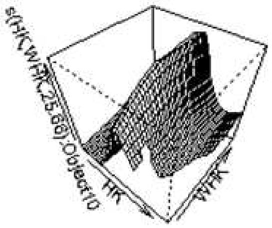


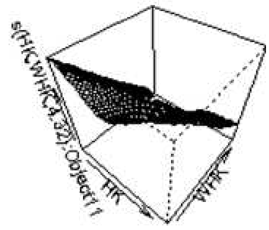




6.A R&D and HK Non-Obj.1 6.B R&D and HK in Obj. 1 6.C R&D and WR&D in Non-Obj.1







6.D R&D and WR&D in Obj.1 6.E HK and WHK in Non-Obj.1 6.F HK and WHK in Obj.1

FIGURE 6

Figure 6 shows the results for the interaction terms for the two sub-groups of regions (same structure as in figure 4 but for the two groups separately). In Graphs 6.A and 6.B R&D is interacted with Human Capital (HK) for Non-Objective 1 and Objective 1 regions respectively. The results show a strong complementarity between Human Capital and R&D investments in Objective 1 regions: the highest the level of human capital in Objective 1 regions the highest the 'productivity' of R&D investments. In Graphs 6.C and 6.D it is R&D activity in neighboring regions (WR&D) to be interacted with internal R&D expenditure in the two sub-groups of regions.

While full-sample results are (again) generally confirmed for non-Objective 1 regions (6.C), for Objective 1 regions it is clear that the capability to benefit from spillovers crucially depends on the internal level of R&D dynamism that ensures the necessary absorptive capacity. Finally, graphs 6.E and 6.F show the results for the interaction between internal (HK) and neighboring regions' Human Capital (WHK). The result for Objective 1 regions is extremely clear cut: we observe a strong shadow effect. When surrounded by other regions with a strong endowment of human capital, lagging regions suffer from the outflow of skilled labour that sharply curbs the impact of internal human capital on innovation. When interpreted in a more systematic fashion these results suggest that Objective 1 regions should invest simultaneously in Human Capital and R&D if they want on the one hand to maximize the complementarities between these two factors of knowledge production and, on the other hand, if they want to exploit external spillovers and minimise outflows of skilled people by improving the matching between their skills profile and local demand.

4 Conclusions

The paper has looked at the genesis of innovation in the EU regions in order to shed new light on the link between innovative inputs (R&D, Human Capital) and the genesis of economically valuable knowledge. The 'traditional' regional Knowledge Production Function (KPF) has been innovatively developed, showing that some relevant innovation dynamics might remain hidden due to some customary - but normally unchallenged - assumptions on its specification and shape. While the positive and significant impact of human capital is easier to capture, the effect of R&D expenditure only emerges after fully controlling for time-varying unobserved heterogeneity (i.e. for a complex set of directly unmeasurable localised socio-institutional conditions that affect the process of innovation). In addition, semi-parametric estimation techniques, that make it possible to relax restrictive assumptions on the shape of the KPF, have uncovered relevant threshold effects for R&D and, to a lesser extent, for Human Capital. A minimum critical mass of these factors needs to be available locally in order to generate innovation. However, the empirical results also suggest a strong complementarity between R&D and Human Capital: both factors need to be targeted simultaneously in order to promote innovation. Furthermore, the econometric analysis has shed light on the relevance of inter-regional spillovers from both R&D and Human Capital. However external factors play a very complex role in the generation of indigenous innovation. The impact of both R&D and Human Capital in neighboring regions crucially depends upon internal conditions that, on the one hand ensure adequate absorptive capabilities and, on the other, prevent/compensate for potential adverse 'shadow effects' from a stronger neighborhood. Finally, even if the dynamics discussed so far apply to the entire EU territory, some of these effects differ between 'core' and 'peripheral' regions. In particular, when restricting the focus on the most disadvantage regions of the European Union (Objective 1 regions) only Human Capital remains the key driver for innovation: Investing in Human Capital in Objective 1 regions is beneficial to innovation irrespective of the their initial endowment and this effect is relatively linear. However, complementary investments in R&D in lagging regions are still necessary in order to counterbalance the absorption of local skills in favour of 'core' areas.

These results have significant implications for innovation policies and the assessment of their spatial impacts. While the Lisbon Agenda placed a significant emphasis on R&D, the increasing EU policy attention for other complementary conditions is fully supported by the evidence produced in this paper. Innovation policies targeting lagging regions should devote particular attention to the assessment of both internal and external conditions in order to design appropriate tailor-made tools to counterbalance potential side-effects of innovation policies. Human Capital and skills should remain the core priority for innovation policies in all contexts.

Even if the results presented in the paper are fully robust, it is important to bear in mind some of the key limitations of the KPF approach. The use of patent intensity as our measure of regional innovation – as customary in the KPF literature - is not problem-free. Different industries and/or different technology sectors tend to show differentiated patenting patterns that are not fully captured by our 'aggregate' analysis. In addition, patents measure invention and completely overlook other forms of innovation (incremental innovation and/or process innovation), potentially under-estimating the innovative potential of more 'disadvantaged' regions.

Further exploration of these dynamics by means of firm-level micro-data is in our agenda for future research.

References

[1] Acs, Z.J., Audretsch D.B., Feldman, M.P. (1994), Real effects of academic research, American Economic Review 82, 363-367.

[2] Anselin L., Acs Z.J., Varga A. (1997), Local geographic spillovers between university research and high technology innovations, Journal of Urban Economics 42, 422-448.

[3] Audretsch D.B. and Feldman M. (2004) Knowledge Spillovers and the Geography of Innovation, in Henderson J.V. and J.F. Thisse (eds.) Handbook of Urban and Regional Economics, Vol.4, pp.2713-2739. Elsevier, Amsterdam

[4] Azomahou, T., T. Mishra, 2008. Age dynamics and economic growth: Revisiting the nexus in a nonparametric setting, Economics Letters vol. 99(1), pages 67-71.

[5] Barca F. (2009) An Agenda for a Reformed Cohesion Policy, European Commission: Brussels

[6] Benavente, J. M. 2006. The role of research and innovation in promoting productivity in Chile, Economics of Innovation and New Technology, 15, 301–315.

[7] Blanchard, P, J.P. Huiban, A. Musolesi, P. Sevestre , 2009, "Where there is a will, there is a way? Assessing the impact of obstacles to innovation?", mimeo.

[8] Bode E. (2004), The spatial pattern of Localized R&D spillovers and empirical investigation for Germany, Journal of Economic Geography, 4, 43-64.

[9] Bottazzi, L., Peri, G. (2003) Innovation, Demand, and Knowledge Spillovers: Evidence from European Patent Data. European Economic Review 47, pp.687–710.

[10] Canning, D. and Pedroni, P. (2004). "The Effect of Infrastructure on Long-run Economic Growth". Harvard University, mimeo.

[11] Cohen, W. M. and D. A. Levinthal, 1989. Innovation and Learning: The two faces of R&D. The Economic Journal 99: 569-596.

[12] Crescenzi, R., Rodríguez-Pose A. and Storper M. (2012) "The territorial dynamics of innovation in China and India", The Journal of Economic Geography, in press.

[13] Crescenzi, R., Rodríguez-Pose A. and Storper M. (2007) "The territorial dynamics of innovation: a Europe-United States Comparative Analysis", The Journal of Economic Geography, 7, 6: 673-709.

[14] Crépon, B., Duguet, E. et Mairesse, J., 1998. Research Investment, Innovation and Productivity: an Econometric Analysis. Economics of Innovation and New Technology 7, 115-158.

[15] Cooke, P. (1998), "Origins of the concept" in Braczyk, H., Cooke, P., and Heidenreich, M. (eds), Regional Innovation Systems. London: UCL Press.

[16] Driscoll J. C., Kraay A. C. (1998), Consistent covariance matrix estimation with spatially dependent panel data, The Review of Economics and Statistics, 80, 549–560.Hastie, T. Tibshirani R. (1990), Generalized Additive Models, Chapman and Hall.

[17] Duguet, E., 2006. Innovation Height, Spillovers and TFP growth at the Firm Level: Evidence From French Manufacturing. Economics of Innovation and New Technology 15, 415-442.

[18] Foddi M., Marrocu E., Paci R. and Usai S. (2012) Knowledge, human capital and regional performance, KIT- ESPON Final Report.

[19] Friedberg, L. (1998), Did Unilateral Divorce Raise Divorce Rates?, American Economic Review 88, 608-627.

[20] Greunz L. (2003) Geographically and Technologically Mediated Knowledge Spillovers between European Regions, Annals of Regional Science 37, 657-80.

[21] Griffith, R., Huergo, E. , Mairesse, J. and B. Peters, 2006. Innovation and productivity across four euroopean countries, Oxford Review of Economic Policy 22(4):,483-498.

[22] Griliches, Z., 1979. Issues in Assessing the Contribution of R&D to Productivity Growth. The Bell Journal of Economics, 92-116.

[23] Hall, Bronwyn, Jacques Mairesse and Pierre Mohnen (2009): 'Measuring the Returns to R&D', CIRANO WP.

[24] Heckman, J.J., Hotz V.J. (1989), Choosing Among Alternative Non-experimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training, Journal of The American Statistical Association, 84(408), 862-874.

[25] Heckman, J.J., E. Vytlacil, 1998. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is Correlated with Schooling," Journal of Human Resources, University of Wisconsin Press, vol. 33(4), pages 974-987.

[26] Henderson, D. J., R. J. Carroll, and Q. Li. 2008. Nonparametric estimation and testing of fixed effects panel data models. Journal of Econometrics 144: 257-275.

[27] Hoderlein, S. and H. White. 2009. Nonparametric identification in nonseparable panel data models with generalized fixed effects. Working Paper, Dept. of Economics, Brown University.

[28] Hsiao C, (2003), Analysis of Panel Data, 2nd edition, Cambridge University Press.

[29] Hsiao C, (2011), "Panel versus Cross-Sectional Data", Plenary lecture, 17th International Panel Data Conference, Montreal.

[30] Huiban J.P. and A. Musolesi (2010) "Innovation and productivity in Knowledge Intensive Business Services", Journal of Productivity Analysis, 1, 63-81.

[31] Iammarino S. (2005) An evolutionary Integrated View of Regional Systems of innovation: concepts, measures and historical perspectives, European Planning Studies 13(4), 497-519.

[32] Jaffe A.B. (1989), Real effects of academic research, American Economic Review 79(5), 957-970.

[33] Lööf, H. and Heshmati, A. 2002. Knowledge Capital and Performance Heterogeneity: a Firm-Level Innovation Study. International Journal of Production Economics, 761. 61–85.

[34] Lööf, H., A. Heshmati, R. Asplund and SO Nås, 2003. Innovation and performance in manufacturing industries: A comparison of the Nordic countries, International Journal of Management Research, 2, 5-36.

[35] Lööf H., Heshmati A., 2006. On the relationship between innovation and performance: a sensitivity analysis. Economics of Innovation and New Technology 15, 317-344.

[36] Lundvall, B.Å. (2001), 'Innovation Policy in the Globalising Learning Economy' in Archibugi, D. and Lundvall, B.Å. (eds.), The globalising learning economy. Oxford University Press.

[37] Mairesse, J. and B. Mulkay, 2007. "An Exploration of Local R&D Spillovers in France," Annales d'Economie et de Statistique, ADRES, issue 87-88, pages 08, Juillet-D.

[38] Mairesse, J. and S. Robin. 2008. Innovation and productivity: a firm-level analysisfor French Manufacturing and Services using CIS3 and CIS4 data (1998-2000 and 2002-2004), Unpublished paper.

[39] Mairesse, J., Mohnen, P., 2005. The importance of R&D for innovation: a reassessment using French survey data, Journal of Technology Transfer, 30(1-2), 183-197.

[40] Mammen, E., Stove, B., and Tjostheim, D., 2009, Nonparametric additive models for panels of time series, Econometric Theory, 25, 442-481.

[41] Masso, J. and P. Vahter. 2008. Technological Innovation and Productivity in Late-Transition Estonia: Econometric Evidence from Innovation Surveys. European Journal of Development Research, Jun2008, Vol. 20 Issue 2, p240-261.Minisian, J. 1962) Research and Development, Production Functions and Rates of Returns. American Economic Review, 492. 80–85.

[42] Midelfart-Knarvik H. and Overman H.G. (2002) Delocation and European integration: is structural spending justified?, Economic Policy 17(35), 322-359

[43] Mohnen, P., Mairesse, J., 2002, Accounting for Innovation and Measuring Innovativeness : An Illustrative Framework and an Application, American Economic Review, Papers and Proceedings, 92(2), 226-230, 2002.

[44] Morgan, K. (1997). "The learning region: Institutions, innovation and regional renewal", Regional Studies, 31 5: 491-503.

[45] Moreno R., R. Paci, S. Usai, (2005) Spatial spillovers and innovation activity in European regions, Environment and Planning A 37, 1793-1812..

[46] Ordas Criado, C., Valente, S. and Stengos, T. (2009), Growth and the pollution convergence hypothesis: A nonparametric approach, MPRA paper.

[47] Overman, H., Redding, S. and Venables, A.J. (2001) 'The economic geography of Trade, Production and Income: a survey of empirics', Centre for Economic Performance, London School of Economics and Political Science

[48] Overman, H. G.; Redding, S.& Venables, A.(2003). The economic geography of trade, production and income : a survey of empirics [online]. London: LSE Research Online. Available at: http://eprints.lse.ac.uk/archive/00000679

[49] O hUallachain B. and Leslie T. F. (2007), Rethinking the regional knowledge production function; *Journal of Economic Geography*, 7, 737-752.

[50] Parisi, M.L., Schiantarelli, F., Sembenelli, A., 2006. Productivity, Innovation and R&D: Micro Evidence for Italy. European Economic Review 50, 2037-2061.

[51] Papke, L.E. (1994), Tax Policy and Urban Development: Evidence from the Indiana Enterprise Zone Program, Journal of Public Economics 54, 37-49

[52] Papke, L.E. (1994), Tax Policy and Urban Development: Evidence from the Indiana Enterprise Zone Program, Journal of Public Economics 54, 37-49.

[53] Ponds R., van Oort F. and Frenken K. (2010), Innovation, Spillovers and university-industry collaboration: an extended knowledge production function approach, Journal of Economic Geography, 10, 231-255.

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

[55] Ruppert, D., Wand, M. P., and Carroll, R. J., 2003, Semiparametric Regression. Cambridge: Cambridge University Press.

[56] Sonn, J.W. and Storper, M. (2008). "The increasing importance of geographical proximity in technological innovation: an analysis of U.S. patent citations, 1975-1997". Environment and Planning A, 40 (5) 1020 - 1039.

[57] Su, L., and A. Ullah. 2006. Profile likelihood estimation of partially linear panel data models with fixed effects, Economics Letters 92: 75-81.

[58] Su, L., and A. Ullah. 2010, Nonparametric and Semiparametric Panel Econometric Models: Estimation and Testing, mimeo.

[59] Tsukay M., Ejiri R., Kobayashi K. and Okumura M. (2007) Productivities of Infrastructure with Spillover Effects: A study of Japan in Karlsson C., Anderson W.P., Johansson B. and Kobayashi K. (eds.) "The management and measurement of infrastructure: performance, efficiency and innovation", Edward Elgar Publishing.

[60] Wood, S.N. (2003) Thin plate regression splines. J.R. Statist.Soc.B 65(1), 95-114.

[61] Wood, S.N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models, Journal of the American Statistical Association. 99:673-686.

[62] Wood S.N. (2006a), Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC Press.

[63] Wood, S.N., (2006b), On confidence intervals for generalized additive models based on penalized regression splines, Australian and New Zealand Journal of Statistics. 48(4): 445-464.

[64] Wood, S.N. (2008), Fast stable direct fitting and smoothness selection for generalized additive models. J.R.Statist.Soc.B 70(3):495-518

[65] Wooldridge J.M, (2005), Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models," The Review of Economics and Statistics, MIT Press, vol. 87(2), pages 385-390.

[66] Wooldridge J.M, (2010), "Econometric Analysis of Cross Section and Panel Data," MIT Press Books, The MIT Press, edition 2,

Table 1: Cobb Douglas function

	(1) Patents p.c.	(2) Patents p.c.	(3) Patents p.c.	(4) Patents p.c
R&D Expenditure	0.260	0.00481	0.600***	0.5834**
	(0.183)	(0.206)	(0.208)	(0.201)
Human Capital	1.397***	0.0382	0.508***	0.529**
	(0.195)	(0.204)	(0.162)	(0.171)
Spillovers from External R&D				2.740
				(1.858)
Spillovers from External HC				-0.855
				(1.327)
One way fixed effects	X	X	X	X
Two way fixed effects		X	X	X
Random growth			X	X
N	1638	1638	1638	1632
R^2	0.918	0.933	0.959	0.9588

All variables are log transformed.*, **, *** significant at 1, 5 and 10 %.
Robust standard errors in brackets.

Appendix 1

Table A.1: Smooth terms and approximate significance

	Smooth terms (EDF)	P. value
R&D Expenditure	7.662	1.71e-12 ***
Human Capital	4.757	1.02e-08 ***
External R&D	3.502	0.315
External HC	7.818	1.08e-12 ***
R&D . Human Capital	28.84	7.7e-11 ***
R&D . External R&D	26.77	< 2e-16 ***
HC . External HC	10.70	0.0567 *
R&D Expenditure non Objective 1	7.266	03e-11 ***
R&D Expenditure Objective 1	1	0.28470
Human Capital non Objective 1	4.636	1.66e-08 ***
Human Capital Objective 1	1.871	0.160472
External R&D non Objective 1	3.162	0.237845
External R&D Objective 1	1	0.779611
External HC non Objective 1	7.969	8.19e-11 ***
External HC Objective 1	3.641	0.000204 ***
R&D . Human Capital non Objective	27.011	3.8e-12 ***
R&D . Human Capital Objective 1	2	0.0451 **
R&D . External R&D non Objective	26.666	< 2e-16 ***
R&D . External R&D Objective 1	3.806	0.7217
HC . External HC non Objective	25.665	< 2e-16 ***
HC . External HC Objective 1	4.319	0.0660 *

*, **, *** significant at 1, 5 and 10 %.

The EDF measure the degree of nonlinearity of the estimated function and when it is equal to 1, this corresponds to a linear relationship.

Appendix 2

We chose a classical and precise geographical definition of neighbourhood based on the Euclidean distance between regions. This scheme imposes smooth distance decay, with weights w_{rj} given by $1/d_{rj}$ where d_{rj} is the Euclidian distance between region r and j for $j \neq r$. The weight matrix (W) is standardized so that the elements in each row sum to 1. R&D spending and Human Capital spillovers are therefore respectively equal to:

$$WRD_{r,t} = \sum_j 1/d_{rj} \times RD_{j,t} \quad \text{and} \quad WHK_{r,t} = \sum_j 1/d_{rj} \times HK_{j,t}$$