SMART SPECIALISATION IN EU REGIONS: REVISITING THE EFFECT OF RELATEDNESS ON REGIONAL PERFORMANCE

Silvia Rocchetta  
University College Dublin  
Spatial Dynamics Lab  
silvia.rochetta@ucd.ie

Dieter Franz Kogler  
University College Dublin  
School of Geography, Planning & Environmental Policy  
dieter.kogler@ucd.ie

Raquel Ortega-Argilés  
University of Birmingham  
Strategy and International Business  
r.ortegaargiles@bham.ac.uk

Abstract
This paper explores the relationship between a region’s knowledge space structure and its labour productivity. To identify the determinants of regional productivity, we perform panel analyses of EU 28 plus Norway NUTS II regions with data covering the 2004-2015 period. We focus on the relationship between related variety and regional labour productivity growth, using the recently developed co-occurrence-based measure of relatedness. This allows us to jointly test the effects of technological variety and relatedness on regional performances. The evidence indicates that both technological relatedness and variety have a non-linear effect on regional productivity. These findings deepen our understanding of the evolution of regional economies and have relevant implications for the implementation of place-specific Smart Specialisation Strategies.
1. Introduction

In the Evolutionary Economic Geography literature there is broad agreement on the positive role of related variety in fostering regional growth. (Frenken and Boschma 2007; Foray, 2014; Cortinovis and Van Oort, 2015). Several empirical studies have provided supporting evidence for the claim that related variety fosters growth in regional employment (Content and Frenken 2016). However, the many empirical studies that tested the effect of related variety on regional productivity have provided results that are often inconsistent with one another.

So far, however, the majority of studies in this stream of the literature have employed the measure of related variety based on the hierarchical structure of official industry classifications (NACE). The fundamental weakness of this approach is that by assuming cognitive similarity to exist only within the same group of the industry classification it fails at capturing the whole range of possibilities by which industries could be related (Bishop and Gripaios 2010; Firgo and Mayerhofer 2017). Moreover, these measures do not allow the possibility of jointly testing the effect of variety and relatedness on regional performance.

On top of this to the best of our knowledge, in the recent literature relatively little attention has been paid so far to the fact that the structure of regional knowledge bases may exert significant influence on regional productivity. In this paper we are specifically interested in analysing how the production and composition (structure) of knowledge production in a given place influences its labour productivity. In particular, our analysis, aims to shed some light on the characteristics of the relationship between the different forms of technological diversification and regional productivity.

Toward this end, we combine information on employment and industrial structure for EU28 countries NUTS II regions with information on patent records. The data sources are Eurostat, Cambridge Econometrics and the European Patent Office’s PATSTAT database.
Following prior art (c.f. Castellani and Pieri, 2016) we measured labour productivity growth using information on the regional gross valued added and regional employment. In order to analyse the impact of technological diversification and related variety on regional productivity we include in our empirical analyses an Information Entropy Index and a Technological Coherence Index. The Information Entropy Index uses information on inventor location and 3-digit technology codes to explore the extent of regions’ technological diversification. The Regional Technological Coherence index, instead, allows us to evaluate the level of epistemic proximity between the patent classes that make up the regional technological structure. Therefore, it represents a proxy for the degree of revealed relatedness between the knowledge components that make up the regional knowledge base.

In our estimation, while focusing on the role of relatedness and variety on regional productivity growth we also take into account the effect of the region’s employment density, stock of technological capital, level of human capital and population size. Inclusion of these control variables is necessary because previous literature (Cingano and Schivardi, 2004; Dettori et al., 2012) has emphasised that regional productivity is likely to depend on the availability of a skilled labour force and the presence of agglomeration economies.

In the econometric analyses, multilevel modelling techniques are applied in order to estimate the degree to which each geographically determined level – e.g. such as national and regional factors - contributes to the explained and unexplained variation of regional productivity. Our hypothesis is that regional productivity is susceptible to both regional and extra-regional contexts.

The produced evidence clearly indicates that both relatedness and variety influence labour productivity growth. Results also highlight that the relationship between technological coherence and regional productivity is non-linear. The results of the analyses confirm our hypothesis that higher levels of technological coherence
may be negatively related with higher levels of regional productivity as they could be linked with potential risks of industrial lock in. While in the case of variety, we found that higher regional performance can only be found at higher levels of industrial technological relatedness.

From a policy perspective, the question of what regional technological profile might make a region more productive has general relevance for the adoption of ‘Smart Specialisation’ Strategies in Europe (McCann and Argilés, 2015). The implementation of Smart Specialization Strategy (S3) involves regional institutions identifying the regional technological specialization domains in which they are expected to concentrate the allocation of resources and which should be the targets of their innovation policy (D’adda et al 2018; McCann & Ortega-Argilés, 2013).

This paper is structured as follows. In section 2 we review the literature on the characteristics of the relationship between relatedness, variety and regional productivity and profile the theoretical line of the paper. In Section 3 we present the data and the variables included in the modelling exercise. In Section 4 we present the estimation strategy, results and the robustness checks. We discuss our main findings and then conclude in Section 5 with reference to relevant implications for the design of regional development policy.

2. Overview and Hypotheses

*Related variety and regional productivity*

Recent enquiries in the Economic Geography literature highlight the positive role of related variety in shaping economic growth. In particular, since the seminal work of Frenken et al. (2007), a number of studies have explored the effect of different measures of related variety on economic performance using different units of analysis. Meanwhile, there’s a common understanding that related variety has a positive impact on regional employment
growth (Frenken et al 2007; Boschma and Iammarino, 2009; Mameli et al., 2012; Cortinovis and Van Oort, 2015).

When it comes to the effect of related variety on regional productivity, however, most studies have provided contrasting results.

Some authors highlight the positive role of related variety on regional productivity. In particular, Falcioğlu (2011) who analysed the determinants of productivity growth in Turkish regions, indicates that related variety has a positive effect on regions’ productivity. Quatraro (2010)’s study corroborates these findings by exploring the evolution of Italian regional economies.

In contrast some studies, provide mixed evidence about the effect of related variety on regional productivity. Using Italian regional trade data, the empirical analyses conducted by Boschma and Iammarino (2009), show that related variety has a positive and significant effect on regional labour-productivity growth in only one econometrical specification out of three. In the same fashion, Bosma et al. (2011), exploiting a dataset on 40 regions in the Netherlands over the period 1988–2002, find that related variety is beneficial for productivity growth in manufacturing sectors only. Opposite results are uncovered for service sectors. More recently Van Oort et al (2015), test the relationship between related variety and productivity growth by conducting an empirical analysis on 205 European NUTS II regions in 15 EU countries between 2000 and 2010. Again, their results were mixed. They highlight, indeed, that related variety does not have any effect on productivity growth in the overall sample, but it has a weak negative effect for small and medium sized urban areas.

The evidence produced by Frenken et al. (2007), using data from 40 NUTS III level Dutch regions, indicates, instead, that related variety has a negative effect on regional productivity growth.
Despite growing scholarly interest, the characteristics concerning the relationship between related variety and productivity are still rather unclear. In a moment in which European regional economies are setting up smart specialisation strategies this ambiguity needs to be resolved. The smart specialisation concept, indeed, originated in the literature in order to explore the reasons of the productivity gap between the United States and Europe (Ortega-Argilés, 2012).

**Characteristics of the relationship between Relatedness and regional productivity**

Since the seminal work by Hildago et al. (2007) several scholars started to calculate relatedness based on co-occurrence analyses (among others: Teece et al 1994; Hidalgo, 2007: Neffke and Henning, 2011; Kogler, 2017). These measures of revealed relatedness allow researchers to better grasp the degree of complementarity or proximity between the components that make up the regional technological structure. In recent Evolutionary Economic Geography literature, there is a common understanding that different forms of complementarities or similarities play a positive role in shaping the opportunities for knowledge recombination (Neffke et al., 2011; Caragliu and Nijkamp, 2015), thus influence the degree of regions’ innovativeness (Kogler, 2013). Recent studies (Boschma et al 2012; Rocchetta and Mina, 2019) indicate that the regional economies’ ability to grow is stronger when their technological structure exhibit a higher degree of revealed relatedness. This implies that regions with a knowledge space that is characterised by a more complementary set of components are better able to maintain a high economic performance. This is likely to be due to the fact that regions that are characterised by a more coherent set of technological classes are more to exploit economies of scope (Kogler et al., 2017).

In the long run, however, a too-strong overlap in competences may lead regional economies to be locked into too specialized technological or industrial paths. An overspecialization in the knowledge space makes regions unable to adapt to the incessantly
changing market requirements (Martin and Sunley, 2006). This was already highlighted by Nooteboom (2000) who claims that there is a trade-off between variety and similarity: too much similarity may result in cognitive lock-in while too little similarity may impede knowledge exchange altogether.

We conjecture therefore that in the short run investing in a coherent set of technologies is beneficial for regional productivity growth. In the long run, however, regional economies need to include in their technological space less similar knowledge domains to maintain satisfactory performances. So far, however, to the best of our knowledge no studies analysed the characteristics of the relationship between related variety and regional productivity using co-occurrence measure of relatedness:

This leads us to formulate our first research question:

**Hypothesis 1: There is a non-linear relationship between regional productivity and the degree of regional technological relatedness**

Several authors suggest variety as one of the major sources of regional economies adaptability (Grabher and Stark, 1997). Regions with a more diversified technological base, indeed are less likely to be locked into declining activities. Following this line of inquiry, Pasinetti (1993) points out that an economy must increase variety in the long run in order to generate productivity gains and limit structural unemployment due to a combination of product innovation and technical progress in production.

Heterogeneity, however, is considered both a cause and an effect of the uncertainty that drives capitalist competition (Schumpeter, 1942; Dosi, 1988). This is also true in the technological space. Recombining more heterogenous pieces of knowledge may indeed be associated with higher risk in the innovation outcome. However, producing innovations based on less related knowledge components, even if it is riskier, might be associated with higher returns in terms of regional economic growth (Fleming, 2001).
Following this line of enquire, we conjecture that in the short-run a too diversified knowledge base will not be beneficial for regional economic growth. In the long run instead, variety is needed to be competitive in the always evolving market scenarios. However, only a few attempts have been made to analyze this trade off by exploring the evolution of the knowledge space. As the effect of variety on the regional economy productivity can be concurrently negative and positive, we formulate the second research question that we will address in this paper:

**Hypothesis 2: There is a non-linear relationship between regional productivity and the degree of regional technological variety**

**Figure 1 – Hypotheses about the relationship between related variety/variety and regional productivity**

The concept of related variety was firstly introduced in the literature in an attempt to resolve the commonly referred to as ‘MAR versus Jacobs’ controversy’. This debate followed an earlier empirical question put forward by Glaeser et al (1992) whether regional economies benefit most from being specialized or being diversified. With MAR externalities they refer to the idea that firm benefit from being located near other companies pertaining to the same industry. This idea was firstly introduced by Alfred Marshall in his seminal work *Principles of economics* (1920), where he claims the existence of knowledge spillovers between firms of
the same sector located in the same place. Jane Jacobs (1969), instead, contended that it is the localized exchange of knowledge among firms pertaining to a diverse set sectors that is more beneficial to innovation and economic growth.

However, as reviewed by Beaudry and Schiffauerova (2009) and De Groot et al. (2016) the many empirical studies on MAR versus Jacobs, have provided results that are often inconsistent with one another. This is probably due to fact that these two types of externalities are not mutually exclusive. It is conceivable that knowledge spillovers between similar industries and across more diverse sectors occur at the same time in the same place. Economic geographers have overlooked the reality that the fastest growing cities and regions in the world, such as New York, London and Shanghai, are usually characterized by the presence of both MAR and Jacobs externalities (Bathelt & Zhao, 2016). The industrial structure of these regions is often marked by the presence of multiple clusters of firms specialized in different industries.

To best of our knowledge however only a very limited number of empirical studies analysed the joint effect of relatedness and variety on regional performance. In this paper we test it by employing a continuous measurement of relatedness that doesn’t force us to strictly classify the technological profile of the region into the specialized or diversified category. The co-occurrence-based measure of relatedness allows us to conceive diversity/specialization as a continuum along a spectrum.

This paper, therefore, aims at contributing to the evolutionary economic geography literature by exploring the relationship between the different form of technological diversification and regional productivity growth.
3. Data and Variables

Dataset

The sample involves 268 NUTS II (Nomenclature of Territorial Units for Statistics) regions in 29 countries (European Union 28 plus Norway). Data concerning regions’ employment and industrial structure were collected from Eurostat and Cambridge Econometrics (ERD). ERD is a service provided by Cambridge Econometrics that provides information on regional gross value added (GVA), gross domestic product (GDP), employment and population. EUROSTAT regional statistics, instead, contains information on workers education level.

Information on patents were extracted from PATSTAT database. We consider patent applications submitted to the EPO (European Patent Office) by inventors residing in one of the regions included in our dataset at the time that the new product or process was produced. Fractional counting is applied in case of multiple inventors per patent coming from different regions. Patent documents are categorised into one or more technology classes using Cooperative Patent Classification (CPC) systems. For the purpose of the present study we analyze patent classes up to 3-digit level. The final panel dataset covers the period 2004 to 2015.

Variables and Measures

Dependent Variable

Following prior art (among others: Castellani and Pieri, 2016) we measured labour productivity growth using the information on the regional gross valued added and regional employment (in thousands of employees) contained in the Cambridge Econometrics database.

The dependent variable, therefore, is calculated over the period 2004-2015 as a log differential between the employment level (GVA) in year t and t-1 divided by employment in the region (pop):

\[
\text{Log differential} = \ln(GVA_t) - \ln(GVA_{t-1}) - \ln(pop_t) + \ln(pop_{t-1})
\]
\[ g_{Productivity}_{kt} = \log \left( \frac{GVA_{k,t}}{Emp_{k,t}} \right) - \log \left( \frac{GVA_{k,t-1}}{Emp_{k,t-1}} \right) \]

**Independent Variables**

As we want to investigate the effect of different degree of technological diversification on regional productivity growth, we introduce in the estimations the variables: Regional Entropy and Coherence. The first one aims at measuring the degree of regional technological diversification. The second instead captures the degree to which the different classes of patents making up the technological knowledge base of a region are related to one other.

The Information Entropy Index was firstly introduced in the economic analysis by Theil in the 1967. In its earliest applications it was used to explore the degree of disorder or randomness of a system (Theil, 1967). Since then has been extensively used in the economics and evolutionary economics geography literature to analyse how different economics activities are distributed among firms, sectors, regions (Attaran, 1986; Frenken et al., 2007; Boschma and Iammarino, 2009). In this paper it aims at measuring the degree of patents portfolio diversification within each region. It captures the degree of regional technological variety.

We exploit patent data at the three-digit level available for each of 268 NUTS II units included in our database to compute Entropy. One of the main peculiarities of this specific measure is that is obtained summing up two different form of variety: related and unrelated. Related variety captures the average degree of disorder or variety within the patent classes and therefore is measured at a lower level of aggregation (3-digit class within a 1-digit section). Unrelated variety, instead, aims at capturing the degree of randomness between the patent classes therefore is computed using higher order patent classes (1-digit section).

More formally we define \( I = (1, \ldots, n) \) as the set containing all 121 patent CPC
subclasses “i” (3-digit technology code). Additionally, $G = (1, ..., g)$ is the set including the eight technological sections of the CPC standard classification "g" (1-digit technology code). Following the CPC standard classification scheme all the patent subclasses $i = 1, ..., n$ in our database fall exclusively under a unique technological section $S_g$, where $g = 1, ..., G$.

Thus, each of the 121 patent subclasses can be grouped into one of 8 technological sections of the CPC standard classification. Along these lines, if we admit that each subclass "i" pertain exclusively to one technological section "g," then we may compute the probability of one patent having the classification "g" ($P_g$) as the summation of probabilities of all subclasses "i" within "g." Formally:

$$P_g = \sum_{i \in S_g} p_i$$

Where $p_i$ stands for the number of patents in the three-digit subclass $i$ within $g$.

Therefore, Unrelated Variety (UV) or between group entropy is calculated as follow:

$$UV = \sum_{g=1}^{G} P_g \log_2 \left( \frac{1}{P_g} \right)$$

The entropy decomposition theorem specifies that the relation between Unrelated Variety and regional Entropy can be formalised as follow:

$$Entropy = UV + \sum_{g=1}^{G} P_g H_g$$

where Related Variety (RV) or within group entropy represents the second part of the equation:

$$RV = \sum_{g=1}^{G} P_g H_g$$
\[
H_g = \sum_{i \in S_g} \frac{p_i}{P_g} \log \left( \frac{1}{\frac{p_i}{P_g}} \right)
\]

Since Hildalgo (2007) seminal work, the evolutionary economic geography literature has highlighted that calculating relatedness following Theil (1967)’s method does not allow to capture the whole range of possibilities by which technologies could be related (Boschma et al., 2012). Its values, indeed, depend entirely on the structure of the Cooperative Patent Classification system. Therefore, several scholars (among others: Neffke et al., 2011; Kogler, 2017) introduced different methods based on co-occurrence matrices to calculate relatedness. As in this paper we focus on how the technological profile of regions shapes productivity our co-occurrence matrix is constructed using patent data. This allows to capture the degree of technological proximity underlying regional structures following the method indicated by Jaffe (1986) and Breschi et al. (2003). Our measure of revealed relatedness the Technological Coherence index (C) is computed exploiting the information contained in EPO patent application documents. In the estimations It is one of our main independent variables because it allows to capture the average degree of cognitive proximity across the patent classes that make up the regional knowledge base (Nesta and Saviotti, 2005 and 2006;).

Following Teece et al. (1994) we firstly compute the Coherence Index \( \tau_{ij} \) to capture the average degree of epistemic proximity between the patent classes that comprise the technological base of NUTS II regions. Our universe includes 268 NUTS II regions each patenting in the period 2004-2015 in some of the 127 technological sections defined by the CPC system. If a NUTS II region \( k \) is producing knowledge in technological class \( i \) then \( i \) \( G_{ik} = 1 \), otherwise \( G_{ik} = 0 \). Accordingly, the total number of regions active in class \( i \) is equal to \( K_i = \sum_k G_{ik} \). In the same fashion, is possible to calculate the total number of regions jointly patenting technology \( i \) and \( j \) : 
\[
O_{ij} = \sum_k G_{ik} G_{jk}.
\]
Applying this formula to all possible pairs of technological classes we compute a square (127 X 127) symmetrical matrix \( \Omega \), in which the
generic cell $O_{ij}$ is the observed number of regions that are jointly patenting in section i and j. To identify cases in which pairs of patent classes are appearing more frequently than randomly would suggest we compare the observed value of $O_{ij}$ with the value that would be expected under the hypothesis that technological diversification is random $\mu_{ij}$ as follow:

$$\tau_{ij} = \frac{O_{ij} - \mu_{ij}}{\sigma_{ij}}$$

where $\mu_{ij}$ is the average of the counterfactual random sample $X_{ij}$,

$$\mu_{ij} = E(X_{ij}) = \frac{K_iK_j}{K}$$

$\sigma_{ij}^2$ is its variance, and

$$\sigma_{ij}^2 = \mu_{ij} \left(1 - \frac{K_i}{K}\right)\left(\frac{K-K_j}{K-1}\right)$$

$K$ is the number of NUTS II regions contained in our database.

Employing the Coherence Index ($\tau_{ij}$) is possible to compute the Weighted Average Relatedness $WAR_{jkt}$ of technology $j$ with respect to all other $m$ technologies present within the region $k$ at time $t$.

$$WAR_{jkt} = \frac{\sum_{m \neq j} \tau_{jm} P_{mk}}{\sum_{m \neq j} P_{mk}}$$

$WAR_{jkt}$ is defined as the degree to which technology $j$ is related to all other technologies $m \neq j$ within the region $k$ (at time $t$), weighted by the number of patent $P_{mk}$ of technology $m$ in the specific NUTS II region at time $t$. Finally, the Regional Technological Coherence (C) of region $k$ at time $t$ is defined as the weighted average of the $WAR_{jkt}$.
\[ C_{kt} = \sum_{j \neq m} \text{WAR}_{jit} \frac{P_{jit}}{\sum_j P_{jit}} \]

where \( \sum_j P_{jkt} \) is the total number of patents within the region \( k \) (NUTS II).

**Control variables**

We include controls for other regional characteristics. We use *Population* to control for size heterogeneity of regional economies and the share of employees with a tertiary level of education (*Education*) for each NUTS II region (that is the number of people between 30-34 with a tertiary degree) to qualify the level of human capital in the region. We expect that a larger share of better educated workforce positively influences regional productivity. Several scholars (among others: Mankiw Et Al., 1992; Benhabib and Spiegel, 1994) underlined the positive role of the level of human capital on productivity.

The regions propensity to produce innovations captured by the variable *Patents*, which is constructed as the yearly stock of patents weighted by population. The literature suggests that the regional propensity to innovate plays a fundamental role in shaping productivity (among others: Dettori et al. 2012). The share of employees in the Knowledge intensive business services (*KIBS*) should be considered to control for the sectorial structure of the region. According to the literature (among others: Corrocher and Cusmano, 2014) Knowledge-intensive business services (*KIBS*) – firms involved in activities such as consultancy, market research, design, engineering and technical services – play a key role in fostering innovation by generating opportunities for interactive learning, favoring the creation of local linkages and contributing to the systems’ connectivity to outside knowledge networks. Taking the ‘quality’ of the industrial structure into account has been underlined as an important determinant of growth differences among regions (among others Paci And Pigliaru, 1999). Following this line of inquiry, we expect that the variable *KIBS* has positive effect on regional productivity growth.
Finally, the regional economies’ size and their degree of wealth is captured by the variable Employment growth density (\textit{gEmployment}). This measure is calculated dividing the employment growth for each NUTS II region by their respective population.

**Descriptive Statistics**

Descriptive statistics and the correlation matrix of dependent and explanatory variables are reported respectively in Table 1 and Table 2.

The correlation matrix shows that there are no variables that are significantly correlated.

**Table 1 - Descriptive statistics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>gLabourproductivity</td>
<td>2948</td>
<td>.007</td>
<td>.046</td>
<td>-.238</td>
<td>.301</td>
</tr>
<tr>
<td>Entropy</td>
<td>2937</td>
<td>3.903</td>
<td>1.511</td>
<td>0</td>
<td>5.894</td>
</tr>
<tr>
<td>Coherence</td>
<td>3148</td>
<td>8.119</td>
<td>1.557</td>
<td>0</td>
<td>12.766</td>
</tr>
<tr>
<td>Education</td>
<td>3086</td>
<td>32.114</td>
<td>11.450</td>
<td>5.6</td>
<td>83.9</td>
</tr>
<tr>
<td>Patents</td>
<td>3066</td>
<td>.119</td>
<td>.167</td>
<td>0</td>
<td>1.131</td>
</tr>
<tr>
<td>gEmployment</td>
<td>2948</td>
<td>.001</td>
<td>.022</td>
<td>-.138</td>
<td>.151</td>
</tr>
<tr>
<td>Population</td>
<td>3216</td>
<td>1812.301</td>
<td>1493.824</td>
<td>26.393</td>
<td>12091.57</td>
</tr>
<tr>
<td>KIBS</td>
<td>3216</td>
<td>.126</td>
<td>.051</td>
<td>.270</td>
<td>.381</td>
</tr>
</tbody>
</table>

**Table 2 - Correlation Matrix**

<table>
<thead>
<tr>
<th></th>
<th>gProd</th>
<th>Entropy</th>
<th>Coherence</th>
<th>Edu</th>
<th>Patents</th>
<th>gEmp</th>
<th>Pop</th>
<th>KIBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>gProd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entropy</td>
<td>-0.067</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>-0.093</td>
<td>0.196</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>-0.032</td>
<td>-0.038</td>
<td>0.001</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patents</td>
<td>-0.085</td>
<td>0.330</td>
<td>0.191</td>
<td>0.136</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gEmp</td>
<td>-0.108</td>
<td>0.012</td>
<td>-0.058</td>
<td>-0.060</td>
<td>0.114</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>0.007</td>
<td>0.0367</td>
<td>0.0290</td>
<td>0.0558</td>
<td>0.0749</td>
<td>-0.0229</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
To gain a first insight about the relationship between regional productivity and relatedness, we plot it in Figure 2. From this graph it seems that Productivity and Relatedness’ have a non-linear relationship and in particular an inverted u-shaped one. This will be considered when designing the estimation strategy.

4. Econometric Analysis and Findings

Several scholars (c.f. Fagerberg and Verspagen, 1996) highlight that productivity is as well shaped by national-specific macroeconomic characteristics such as inflation, public deficit and external accounts. Therefore, in order to take into account both national and regional factors that shape regional productivity growth we approach our empirical analysis
employing and a multilevel model (MLM). Multilevel analysis is a suitable empirical strategy for the treatment of the nested structure of our dataset as it allows to estimate and decompose the total random variation of the regional and country level explanatory variables. The residual variance is, indeed, partitioned into a between-regions component (the variance of the country-level residuals) and a within-countries component (the variance of the regional-level residuals). The main advantage of MLM is that it allows to treat the unit of an analysis as independent. Other methods to analyze clustered data such as cluster-adjusted ordinary, instead, does not account for the hierarchical structures of the data and treat the random variation as ‘disturbance’. This implies that these techniques underestimate higher-level independents variables standard errors’ coefficients.

The estimations use random-intercept models with level-one (regional level) and level-two (national level) covariates:

\[
g\text{Labourproductivity}_{ijt} = \beta_0 + \beta_1 \text{Variety}_{t-1} + \beta_2 \text{Variety}^2_{ijt-1} + \beta_3 \text{Coherence}_{ijt-1} + \\
\beta_4 \text{Coherence}^2_{ijt-1} + \beta_5 \text{Controls}_{ijt-1} + \mu_{i(j)t} \ \nu_{ijt}n
\]

In order to avoid problems of reverse causality all our explanatory variables are lagged by one period.

Table 3 reports the results of the estimations. The results of the estimations (from Column 1 to 3) show that Technological Coherence (Coherence) appear to have a positive and significant effect on productivity growth while squared Technological Coherence is still significant but with a negative coefficient. This evidence confirms our first hypothesis according to which the relationship between Technological coherence and regional productivity growth is non-linear and presents an inverted U-shape. Higher levels of technological coherence may be negatively related to higher levels of regional performance linking with risks of potential technological lock-in.
Our estimations also highlight that entropy has negative and significant effect on productivity growth while squared entropy positively influence regional productivity. This implies that relationship between variety and regional productivity is non-linear and it is U-shaped. These results confirm our second hypotheses.

Following our conjectures, the model yields positive and significant coefficients for the variables education and patents, indicating that regions with higher levels of human capital and more innovative tends to be more productive. It is also worth noting that the variable KIBS, which, shows negative and significant coefficients

Robustness checks
To further test the robustness of the results derived from the initial modeling exercise, the same multilevel estimations are performed using fixed effects (column 3 table 3) and employing a different non-linear transformation of our focal independents variable Entropy and Coherence.

The results of the model that use a cubic transformation of our focal explanatory variable Entropy and Coherence are reported in Table 4. Both analyses corroborate the results derived in the initial estimation.

5. Conclusions
This paper contributes to the conceptual and empirical development in evolutionary economic geography focusing on the role of related technological and industrial variety in shaping regional performances. More specifically, it complements and augments the literature on relatedness and the conceptualization of the role of specialisation and diversification on regional productivity growth.

The present study explored the factors affecting the labour productivity growth of EU28 plus Norway NUTS II regional economies. We conjectured that a fundamental determinant has to reside in the regional knowledge space structures.
Our panel analyses on EU28 plus Norway NUTS II regions highlights that both relatedness and variety affect labour productivity growth. The evidence produced in the study also confirm our hypotheses about the non-linearity of the relationship between the degree of regional technological relatedness/variety and labour productivity growth. In particular, the empirical analyses uncover the fact that the relationship between relatedness and regional productivity growth is non-linear and presents an inverted U-shape. While the one between variety and regional productivity is non-linear and it is U-shaped.

This implies that activities aimed at the technological upgrading of the regional economies must be built around related technologies (Kogler, 2017; Boschma, 2012). At the same time, however, each region, once it has built its portfolio of related technologies needs to diversify its knowledge base to maintain a satisfactory level of regional growth in the long run.

The contribution of this paper is therefore threefold: firstly, the evidence outlined in the present contribution demonstrates that there is a non-linear relationship between regional productivity and technological relatedness/variety; secondly it significantly adds to the existing literature by showing that the degree of both technological relatedness and variety jointly affect regional productivity growth; and thirdly, in the analysis we test the effect of relatedness on regional productivity using a co-occurrence based measure of revealed relatedness between the technologies that make the regional industrial structure.

The study has, of course, limitations. Potential limitations of our analysis are related to the methodology that we use to capture the non-linear effects. Our analysis only considers one inflection point in the distribution of relatedness and variety. The inclusion of a co-occurrence-based measure of technological diversity would be ideal in order to better grasp the level of the regional technological diversification. Moreover, complementary firm-level studies would be ideally placed to dig deeper into this more specific question. The modelling
of knowledge interdependencies might prove important also in this stream of research and could extend further the analysis of productivity in a multi-level (firm-region) framework.

There is a clear connection between the evidence produced in this paper and the implementation of Smart Specialisation strategies. In the last years, many EU regions have been involved in the implementation of Regional Innovation and Smart Specialisation Strategies (RIS3). While many European countries and regions were strong in developing new technologies and techniques in leading technology sectors, they appeared to be systematically much weaker in adopting and adapting these technologies to a wider range of sectors, activities and locations, beyond the new technology sectors themselves. Fostering entrepreneurial actions which are built on technologies, sectors or activities which exhibit both local scale and embeddedness is essential (McCann and Ortega-Argilés 2015). A platform for entrepreneurial and innovation promotion is critical in order for small and incremental innovations to display sufficiently large scale effects to help transform the existing system, and it is imperative that the mobilization of activities, technologies or sectors with potential scale are prioritised. At the same time activities aimed at the technological upgrading and diversification of the system must also be built around the system’s existing capabilities and skill sets, or rather what is known as ‘related variety’ (Frenken et al. 2007; Frenken and Boschma 2007), as this maximizes the chances of long run success and learning. Allied to these dimensions, efforts aimed at promoting knowledge connectivity and knowledge spillovers must operate both at fostering greater local intra-regional linkages as well as wider inter-regional and international knowledge linkages. These latter points are especially important in today’s economy where global value chains have reconfigured numerous commercial and production relationships. Identifying those technologies, activities or sectors which are able to better leverage global value chains is also imperative in order to build scale and connectivity. At the same time, and as would be expected from the smart specialisation principles, we also see significant variations in priority areas or themes.
between different regions (McCann and Ortega-Argilès 2016). While the uptake in Northern European regions has generally been relatively smooth, the policy agenda appears to have been particularly beneficial to many Southern European regions in helping improve and enhance the policy settings and their policy design and delivery processes. In contrast, there are still major challenges in Central and Eastern European regions and member states (McCann and Ortega-Argilès 2016), and improving the policy design and delivery in these localities will continue to be an ongoing priority.
Table 3 – Estimations using Multilevel Models, Multilevel Models with standardized coefficient and Foxed effects.

<table>
<thead>
<tr>
<th></th>
<th>(2) MLM gLabourproductivity</th>
<th>(2) MLM Stand Coeff gLabourproductivity</th>
<th>(3)FE gLabourproductivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entropy</td>
<td>-0.00539**</td>
<td>-0.00830**</td>
<td>-0.00665**</td>
</tr>
<tr>
<td></td>
<td>(0.00205)</td>
<td>(0.00316)</td>
<td>(0.00231)</td>
</tr>
<tr>
<td>SqEntropy</td>
<td>0.000936***</td>
<td>0.00930**</td>
<td>0.00117**</td>
</tr>
<tr>
<td></td>
<td>(0.000325)</td>
<td>(0.00323)</td>
<td>(0.000374)</td>
</tr>
<tr>
<td>Coherence</td>
<td>0.00568**</td>
<td>0.0100**</td>
<td>0.00769**</td>
</tr>
<tr>
<td></td>
<td>(0.00200)</td>
<td>(0.00354)</td>
<td>(0.00239)</td>
</tr>
<tr>
<td>SqCoherence</td>
<td>-0.000429*</td>
<td>-0.00798*</td>
<td>-0.000621**</td>
</tr>
<tr>
<td></td>
<td>(0.000184)</td>
<td>(0.00342)</td>
<td>(0.000215)</td>
</tr>
<tr>
<td>Education</td>
<td>0.000323*</td>
<td>0.00373*</td>
<td>0.000912***</td>
</tr>
<tr>
<td></td>
<td>(0.000133)</td>
<td>(0.00153)</td>
<td>(0.000238)</td>
</tr>
<tr>
<td>Patents</td>
<td>0.0159**</td>
<td>0.0266**</td>
<td>0.153***</td>
</tr>
<tr>
<td></td>
<td>(0.00579)</td>
<td>(0.000971)</td>
<td>(0.0186)</td>
</tr>
<tr>
<td>Gemp</td>
<td>0.0265</td>
<td>0.000598</td>
<td>0.0220</td>
</tr>
<tr>
<td></td>
<td>(0.0362)</td>
<td>(0.000817)</td>
<td>(0.0387)</td>
</tr>
<tr>
<td>Population</td>
<td>0.0000000570</td>
<td>0.000852</td>
<td>0.0000125</td>
</tr>
<tr>
<td></td>
<td>(0.000000548)</td>
<td>(0.000819)</td>
<td>(0.0000157)</td>
</tr>
<tr>
<td>KIBS</td>
<td>-0.0944***</td>
<td>-0.00485***</td>
<td>-0.0348</td>
</tr>
<tr>
<td></td>
<td>(0.0277)</td>
<td>(0.00143)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>Regional Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0140</td>
<td>0.0276***</td>
<td>-0.0278</td>
</tr>
<tr>
<td></td>
<td>(0.00784)</td>
<td>(0.00396)</td>
<td>(0.0194)</td>
</tr>
<tr>
<td>Obs</td>
<td>2948</td>
<td>2948</td>
<td>2948</td>
</tr>
<tr>
<td>N of Regions</td>
<td>268</td>
<td>268</td>
<td>268</td>
</tr>
</tbody>
</table>

Estimated intercept and slope coefficients for each regressor with robust standard errors in parentheses.
Asterisks denote significance: * p<0.05, ** p<0.01, *** p<0.0
Table 4 – Estimations using Multilevel Models, Multilevel Models with standardized coefficient and Foxed effects including cubic entropy and coherence

<table>
<thead>
<tr>
<th></th>
<th>(2) MLM</th>
<th>(2) MLM Stand Coeff</th>
<th>(3)FE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gLabourproductivity</td>
<td>gLabourproductivity</td>
<td>gLabourproductivity</td>
</tr>
<tr>
<td>Entropy</td>
<td>-0.00345**</td>
<td>-0.00532**</td>
<td>-0.00378**</td>
</tr>
<tr>
<td></td>
<td>(0.00131)</td>
<td>(0.00202)</td>
<td>(0.00137)</td>
</tr>
<tr>
<td>CubicEntropy</td>
<td>0.000113**</td>
<td>0.00661**</td>
<td>0.000133***</td>
</tr>
<tr>
<td></td>
<td>(0.000361)</td>
<td>(0.00210)</td>
<td>(0.0000390)</td>
</tr>
<tr>
<td>Coherence</td>
<td>0.00417**</td>
<td>0.00737**</td>
<td>0.00527***</td>
</tr>
<tr>
<td></td>
<td>(0.00130)</td>
<td>(0.00229)</td>
<td>(0.00143)</td>
</tr>
<tr>
<td>CubicCoherence</td>
<td>-0.0000274*</td>
<td>-0.00496*</td>
<td>-0.0000375**</td>
</tr>
<tr>
<td></td>
<td>(0.0000108)</td>
<td>(0.00195)</td>
<td>(0.0000115)</td>
</tr>
<tr>
<td>Education</td>
<td>0.000336*</td>
<td>0.00388*</td>
<td>0.000785***</td>
</tr>
<tr>
<td></td>
<td>(0.000134)</td>
<td>(0.00154)</td>
<td>(0.000218)</td>
</tr>
<tr>
<td>Patents</td>
<td>0.0156**</td>
<td>0.00262**</td>
<td>0.143***</td>
</tr>
<tr>
<td></td>
<td>(0.00579)</td>
<td>(0.000970)</td>
<td>(0.0175)</td>
</tr>
<tr>
<td>Gemp</td>
<td>0.0209</td>
<td>0.000471</td>
<td>0.0109</td>
</tr>
<tr>
<td></td>
<td>(0.0361)</td>
<td>(0.000816)</td>
<td>(0.0387)</td>
</tr>
<tr>
<td>Population</td>
<td>0.000000567</td>
<td>0.000847</td>
<td>0.0000122</td>
</tr>
<tr>
<td></td>
<td>(0.000000548)</td>
<td>(0.000818)</td>
<td>(0.0000127)</td>
</tr>
<tr>
<td>KIBS</td>
<td>-0.0968***</td>
<td>-0.00498***</td>
<td>-0.0300</td>
</tr>
<tr>
<td></td>
<td>(0.0279)</td>
<td>(0.00143)</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Regional Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year Fixed Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0124</td>
<td>0.0280***</td>
<td>-0.0648*</td>
</tr>
<tr>
<td></td>
<td>(0.00789)</td>
<td>(0.00405)</td>
<td>(0.0280)</td>
</tr>
<tr>
<td>Obs</td>
<td>2948</td>
<td>2948</td>
<td>2948</td>
</tr>
<tr>
<td>N of Regions</td>
<td>268</td>
<td>268</td>
<td>268</td>
</tr>
</tbody>
</table>

Estimated intercept and slope coefficients for each regressor with robust standard errors in parentheses. Asterisks denote significance: * p<0.05, ** p<0.01, *** p<0.0
References


Diego D’Adda, Enrico Guzzini, Donato Iacobucci & Roberto Palloni (2018) Is Smart Specialisation Strategy coherent with regional innovative capabilities?, Regional Studies on line only.


