



Co-evolutionary patterns in regional knowledge bases and economic structure: evidence from European Regions

Francesco QUATRARO

University of Nice Sophia Antipolis, GREDEG-CNRS

250, rue Albert Einstein

06560, Valbonne (France)

Tel: +33 4 93954373

Email address: Francesco.quatraro@unice.fr

and

BRICK (Bureau of Research on Innovation, Complexity and Knowledge), Collegio Carlo Alberto

ABSTRACT. This paper presents an analysis of the co-evolutionary patterns of structural change in knowledge and economics. The former is made operational through an analysis of co-occurrences of technological classes in patent documents in order to derive indicators of coherence, variety and cognitive distance. The latter, on the other hand, is made operational in a synthetic way by implementing shift share analysis which decomposes labour productivity growth into effects caused by changes in the allocation of employment, those ascribed to intra-sector productivity growth and those caused by interaction of these two components. The results of the analysis conducted on a sample of 227 European regions show that increasing variety is associated with the reallocation of workforce across sectors whereas within sector productivity is associated with high levels of both coherence and cognitive distance of the regional knowledge base.

JEL Classification Codes: O33, R11

Keywords: Recombinant Knowledge, Coherence, Variety, Regional Structural Change, Shift Share Analysis



1 Introduction¹

The relationships between technological change and structural change have not received due attention in empirical analyses. Indeed, the two subjects have been mostly analysed independently or in relation to the mechanisms of economic growth, but not much consideration has been given to their close connection.

Within the domain of regional economics, an early effort to integrate analysis of technological and structural change was made by François Perroux (1955) who integrated the role of technological change in his “growth pole” theory. Regional economic systems are characterized by rounds of growth, i.e. periods in which firms in the propulsive industry grow at faster rates, propagating positive effects across firms directly and indirectly related to the propulsive industry. The main driving factor in this expansion is technical efficiency achieved through innovation efforts.

More recently, the evolutionary approach to economic geography has articulated a framework for analysis of linkages between technology and structural change at the local level that is able to account for the inherent complexity of the dynamics at stake (Boschma and Frenken, 2006 and 2011). In this strand of analysis, the competences accumulated at the local level are important in shaping the process of industrial diversification so that a change in the allocation of the workforce across different sectors is influenced by the degree of technological relatedness between the involved activities.

This paper aims to contribute to the debate by bringing knowledge into an empirical setting by analysing its relationship with cross-regional differential structural change dynamics. In doing so, we will draw on the recombinant knowledge approach to propose the concept of regional knowledge structure (Weitzman, 1998; Fleming, 2001; Fleming and Sorenson, 2001; Quatraro, 2012). We will then couple three different methodological approaches. The former elaborates a set of indicators that are able to provide a synthetic account of the architecture of knowledge structure. In particular, we will draw upon co-occurrence matrixes to calculate coherence, cognitive distance and variety indicators.

¹ A preliminary version of this paper has been presented at the annual conference of the Italian branch of the Regional Science Association (AISRE) held in Rome in September 2012. The funding of European Commission through the FP7 research project ‘Policy Incentives for the Creation of Knowledge: Methods and Evidence’ (PICK-ME), Grant number 266959, is also gratefully acknowledged.



Secondly, we will provide a synthetic account of the change in economic structure by implementing “shift-share analysis” in order to disentangle the contribution to (labour) productivity growth of within-sector productivity dynamics and reallocation of the labour force across the different sectors. Finally, the relationships between these two sets of indicators will be investigated by using a vector autoregression (VAR) model which we estimate via ‘reduced form’ by applying the least absolute deviation (LAD) estimator due to the distributional properties of the variables.

The analysis is carried out on European NUTS II regions and provides an interesting insight into the dynamic feedbacks between economic and knowledge structure. Innovation patterns dominated by established capabilities and organized search are indeed related to productivity gains in fast growing sectors whereas innovation patterns typical of the early stages of technological trajectories and dominated by random screening are mostly related to changes in the cross-sectoral allocation of employment. The rest of the paper is organized as follows. The next section outlines the theoretical framework, Section 3 elaborates a model, introduces ‘shift-share’ analysis and proposes the working hypotheses and Section 4 provides a description of the data used and outlines the econometric strategy. Section 5 presents the results of the estimations and a general discussion and Section 6 offers a final conclusion.

2 Economic Structure and Knowledge Structure: the missing link

A wide body of literature in the domain of economics of innovation deals with the importance that knowledge plays in the economic development process. On the contrary, scarce attention has been devoted to the relationship between technological knowledge and the dynamics of structural change². This is particularly true for what concerns empirical analyses grounded on econometric modeling. Indeed, most of the studies in this field revert to case studies. For example, Lever (2002) provides an assessment of the link between knowledge base and economic growth in 19 European cities. An important attempt to link regional innovation systems to local knowledge bases can be also found in Asheim and Coenen (2005), wherein the authors show the overlapping of features of the knowledge base and regional industrial specialization by focusing on Nordic clusters.

2



The limited attention to the relationship between technological knowledge and structural change is all the more surprising for at least two reasons. First, the process of economic development is punctuated by changes in the relative weight of sectors (Kuznets, 1930; Burns, 1934; Fisher, 1939). Second, the close relationship between structural change and technological change was already clear to influential scholars such as Kuznets (1930) and Schumpeter (1939 and 1942), who, however, did not elaborate much on this from an analytical viewpoint³. The interplay between Schumpeterian dynamics and the retardation theory provides a fertile ground for improving our understanding of regional differences in the transition dynamics typical of structural change processes (Quatraro, 2009a).

The evolutionary economic geography approach (Boschma and Frenken, 2006 and 2011) proposes a far reaching integration of these issues in a framework explicitly combining industrial dynamics with economic geography. The main argument is that regional growth emerges out of a process of industrial diversification. The emergence of new industries at the local level, i.e. the shift of employment away from one sector to another, is influenced by technological relatedness between sectors. Proximity matters not only in the geographical but also in the technical and technological space so that the introduction of new varieties is constrained by the competencies accumulated at the local level (Boschma, 2005; Usai, 2011; Antonelli and Quatraro, 2013).

While the extant literature has focused on the role of technological relatedness of sectors in the process of regional branching, less attention has been devoted to the importance of the emergence of new sectors of knowledge accumulated at the local level. In this direction, the grafting of the recombinant knowledge approach onto the ongoing debate may enhance understanding of the relationship between patterns of knowledge generation and structural change.

The recombinant knowledge approach provides a far reaching framework for representing the internal structure of regional knowledge bases as well as enquiring into the effects of its evolution. If knowledge stems from a combination of different technologies, the

³ François Perroux (1955) provided the former efforts to build a systematic account of the relationships between structural and technological change in local contexts. He proposed a view according to which the development of local economies is shaped by centripetal and centrifugal forces. Some sectors are likely to be stronger in some areas and weaker in others so that the economic development of a specific area is influenced by the structural ties of the propulsive sector with the rest of the local economic activities. Vertical and horizontal linkages can therefore enhance the positive effects of outperforming sectors. Of the main sources of competitive advantage in this framework, innovation plays a key role in the development of technical efficiency.



frequency with which two technologies are combined provides useful information on the basis of which the internal structure of the knowledge base can be characterized according to the average degree of complementarity of the technologies which knowledge bases are made of as well as the variety of the observed pairs of technologies (Fleming, 2001; Fleming and Sorenson, 2001).

The dynamics of regional knowledge bases can be seen as the patterns of change in their own internal structure, i.e. the patterns of recombination across the elements in the knowledge space (Quatraro, 2010). This qualifies both the cumulative character of knowledge creation and the key role played by the properties describing knowledge structure, as well as linking them to the relative stage of development of a technological trajectory (Dosi, 1982; Saviotti, 2004 and 2007; Krafft, Quatraro, Saviotti, 2014). The ability to engage in a search process within cognitive spaces that are close to the core competencies residing in the region is indeed the outcome of a learning process that displays its effects once the technological trajectory gets well established. On the contrary, in the wake of a new technological trajectory, search processes are more likely to occur in areas that are distant from existing competencies in the cognitive space (Nightingale, 1998; Sorenson et al., 2006).

The intertwining of regional innovation capabilities and recombinant knowledge is of great importance in understanding the relationships between knowledge and structural change. Innovation capabilities are indeed unevenly distributed across regions and technological trajectories so that the emergence of new sectors is more likely to be associated with periods of random screening in the knowledge space, typical of the early phases of new trajectories in which capabilities are not yet sufficiently established (Lawson and Lorenz, 1999; Romijn and Albu, 2002; Todling and Trippl, 2005; Antonelli, 2008; Quatraro, 2009b). As the trajectory gets more familiar, innovating agents learn to move across the knowledge space and are more likely to undertake organized search directed towards the combination of technologies that are close to one another. The transition to organized search is typical of phases in which the recombination activity occurs out of a sharply defined region of the knowledge space. The likelihood of successful innovations is greater in this stage and they are more likely to foster productivity growth at the regional level rather than engender changes in the economic structure (Krafft, Quatraro and Saviotti, 2011 and 2014).

In view of the arguments outlined so far, we are now able to spell out our working hypotheses. Technological change and structural change are closely intertwined. Indeed, the



accumulation of technological knowledge is likely to shape the processes by which new industries, either in relative or absolute terms, emerge in regional contexts. Technological knowledge, however, ought to be regarded as an heterogeneous asset stemming from the combination of a variety of pieces of knowledge. The way these pieces are combined together by regional innovating actors gives the structure of the local knowledge base a peculiar shape. The economic structure can therefore be related to the structure of technological knowledge. The latter may have different configurations according to the search patterns underlying its production which are in turn related to the relative maturity of the technological trajectory they impinge upon. When a technology trajectory is sufficiently established and innovation capabilities diffused at the regional level, search patterns are likely to be dominated by organized efforts directed towards well defined areas of the knowledge space. Innovating agents are more able to identify pieces of knowledge that can be successfully combined together and feature high complementarity and similarity. Knowledge at this stage is likely to have a boosting effect on the productivity dynamics of established sectors. In the early stages of a technological trajectory, innovation capabilities in the new domain still have to be fully developed and diffused so that innovating agents still move in a highly uncertain environment. Search patterns are dominated by random screening in the knowledge space which is likely to lead to the combination of pieces of knowledge loosely related to one another and hence with low levels of coherence and similarity. The switch to a new technological trajectory is likely to be associated with the reallocation of employment across sectors rather than with productivity gains of established sectors.

3 The implementation of key variables

3.1 Shift share decomposition

Shift share analysis provides an interesting methodology that allows labour productivity to be decomposed in order to identify the differential contribution provided by changes in the reallocation of employment across sectors, i.e. the most traditional utilization of the concept of structural change in economics.

As noted by Houston (1967), the origins of shift-share analysis can be dated back to seminal work by Daniel Creamer (1942), although it did not reach great success at least until 1960, when Perloff, Dunn, Lampard and Mutt employed it as an analytical tool in their work, *Regions, Resources and Economic Growth*. It has been mostly used to investigate and disentangle the compositional mix and competitive position of regions in the face of observed changes in some relevant variables (Esteban, 1972 and 2000). In this paper, we will follow the approach developed by Fagerberg (2000) who decomposed labour productivity into three major components, i.e. the allocative and the productivity differential and interaction between the two. We start by rearranging labour productivity as follows (region subscripts are omitted for the sake of clarity):

$$\frac{Y}{L} = \frac{\sum_j Y_j}{\sum_j L_j} = \sum_j \left[\frac{Y_j}{L_j} \frac{L_j}{\sum_j L_j} \right] \quad (1)$$

Labour productivity at the system level can therefore be decomposed into the contribution provided by labour productivity of each sector j as well as by share of sector j in total employment.

If we set:

$$P_j = \frac{Y_j}{L_j} \quad (2)$$

$$S_j = \frac{L_j}{\sum_j L_j} \quad (3)$$

Then:

$$\frac{Y}{L} = \sum_j [P_j S_j] \quad (4)$$

The variation in labour productivity can therefore be expressed as follows:

$$\Delta \frac{Y}{L} = \sum_j [P_{j,t-1} \Delta S_j + \Delta P_j \Delta S_j + S_{j,t-1} \Delta P_j] \quad (5)$$

Equation (5) can therefore be expressed in growth rates by dividing it by (Y/L) :

$$\frac{\Delta(Y/L)}{(Y/L)} = \sum_j \left[\frac{P_{j,t-1} \Delta S_j}{(Y/L)} + \frac{\Delta P_j \Delta S_j}{(Y/L)} + \frac{S_{j,t-1} \Delta P_j}{(Y/L)} \right] \quad (6)$$

The first term between parentheses is the contribution to productivity growth from **changes in the allocation of labour** between industries. It will be positive if the share of high productivity industries in total employment increases at the expense of industries with low productivity. The second term measures the **interaction between changes in productivity** in individual industries **and changes in the allocation of labour** across industries. It will be positive if fast growing sectors in terms of productivity also increase their share in total

employment. The third term is the **contribution from productivity growth within industries**.

For the sake of clarity, let us assign a symbol and a name to each of the identified components:

$$\text{Reallocation term} = \mu = \sum_j \frac{P_{j,t-1} \Delta S_j}{(Y/L)} \quad (7)$$

$$\text{Cross-term} = \pi = \sum_j \frac{\Delta P_j \Delta S_j}{\left(\frac{Y}{L}\right)} \quad (8)$$

$$\text{Within-sector productivity} = \alpha = \sum_j \frac{S_{j,t-1} \Delta P_j}{(Y/L)} \quad (9)$$

3.2 The Knowledge Indicators

The implementation of regional knowledge indicators rests on the recombinant knowledge approach and on the model described in Section 2. In order to provide an operational translation of such variables, we need to identify both a proxy for the pieces of knowledge and a proxy for the elements that form their structure. For example, we could take scientific publications as a proxy for knowledge and look either at keywords or at scientific classification (the JEL code for economists, for example) as a proxy for the constituting elements of the knowledge structure. Alternatively, we could consider patents as a proxy for knowledge and then look at technological classes to which patents are assigned as the constituting elements of its structure, i.e. the nodes of the network representation of recombinant knowledge. In this paper we will follow the latter path⁴. Each technological class j is linked to another class m when the same patent is assigned to both of them. The higher the number of patents jointly assigned to classes j and m , the stronger this link is. Since technological classes attributed to patents are reported in the patent document, we will refer to

⁴ The limits of patent statistics as indicators of technological activities are well known. The main drawbacks can be summarized in their sector-specificity, the existence of non-patentable innovations and the fact that they are not the only protecting tool. Moreover, the propensity to patent tends to vary over time as a function of the cost of patenting and is more likely to feature large firms (Pavitt, 1985; Griliches, 1990). Nevertheless, previous studies highlighted the usefulness of patents as measures of production of new knowledge, above all in the context of analyses of innovation performances at the regional level. Such studies show that patents represent very reliable proxies for knowledge and innovation, as compared with analyses drawing upon surveys directly investigating the dynamics of process and product innovation (Acs et al., 2002). Besides the debate about patents as an output rather than an input of innovation activities, empirical analyses showed that patents and R&D are dominated by a contemporaneous relationship, providing further support to the use of patents as a good proxy of technological activities (Hall et al., 1986). Moreover, it is worth stressing that our analysis focuses on the dynamics of manufacturing sectors.

the link between j and m as the co-occurrence of both of them within the same patent document⁵.

The four properties of the knowledge base (KB) which we will use in our analysis are the knowledge capital stock, its variety, related or unrelated, its coherence and the cognitive distance (see the Appendix for details).

The traditional **regional knowledge stock** (KCAP) is computed by applying the permanent inventory method to patent applications. We calculated it as the cumulated stock of past patent applications using a rate of obsolescence of 15% per annum:

$E_{i,t} = \dot{h}_{i,t} + (1 - \delta)E_{i,t-1}$, where $\dot{h}_{i,t}$ is the flow of regional patent applications and δ is the rate of obsolescence.

The **variety** (KV) of a knowledge base measures the extent of its diversification, with related variety measuring it at a lower level of aggregation and unrelated variety at a higher level of aggregation (Frenken et al, 2007). Technological variety can be measured by using the information entropy index. It was introduced by Shannon (1948) to measure the information content of messages and can be used as a distribution function in a number of circumstances (Theil, 1967, Frenken 2004).

Systems characterized by high entropy will also be characterized by a high degree of uncertainty (Saviotti, 1988). Unlike common measures of variety and concentration, information entropy has some interesting properties (Frenken and Nuvolari, 2004). An important feature of the entropy measure is its multidimensional extension. Consider a pair of events (X_i, Y_j) , and the probability of co-occurrence of both of them p_{ij} . A two dimensional (total) knowledge variety (KV) measure can be expressed as follows:

$$TV \equiv H(X, Y) = \sum_l \sum_j p_{lj} \log_2 \left(\frac{1}{p_{lj}} \right) \quad (10)$$

If we consider p_{lj} to be the probability that two technological classes l and j co-occur within the same patent, then the measure of multidimensional entropy focuses on the variety of co-occurrences of technological classes within regional patents applications.

Moreover, the total index can be decomposed into a “within” and a “between” part anytime the events to be investigated can be aggregated into a smaller numbers of subsets.

⁵ It must be stressed that to compensate for intrinsic volatility of patenting behaviour, each patent application is made to last five years.

Within-entropy measures the average degree of disorder or variety within the subsets, while between-entropy focuses on the subsets measuring the variety across them. Frenken et al. (2007) refer to between- and within-group entropy respectively as unrelated and related variety.

It can be easily shown that the decomposition theorem also holds for the multidimensional case. Hence, if we allow $l \in S_g$ and $j \in S_z$ ($g = 1, \dots, G$; $z = 1, \dots, Z$), we can rewrite $H(X, Y)$ as follows:

$$KV \equiv H(X, Y) = \sum_l \sum_j p_{lj} \log_2 \left(\frac{1}{p_{lj}} \right) \quad (11)$$

where the first term of the right-hand-side is the between-entropy and the second term is the (weighted) within-entropy. In particular:

$$UKV \equiv H_Q = \sum_{g=1}^G \sum_{z=1}^Z P_{gz} \log_2 \frac{1}{P_{gz}} \quad (12)$$

$$RKV \equiv \sum_{g=1}^G \sum_{z=1}^Z P_{gz} H_{gz} \quad (13)$$

$$P_{gz} = \sum_{l \in S_g} \sum_{j \in S_z} p_{lj}$$

$$H_{gz} = \sum_{l \in S_g} \sum_{j \in S_z} p_{lj} \log_2 \left(\frac{1}{p_{lj}} \right)$$

We can therefore refer to between- and within-entropy respectively as *unrelated knowledge variety (UKV)* and *related knowledge variety (RKV)*, while total information entropy is referred to as *general technological variety*.

The **coherence** (COH) of a KB measures the extent to which different types of knowledge can be combined. This is of fundamental importance since the types of knowledge required by innovating agents to create new products or services are not necessarily found within a discipline, but need to be combined to produce the desired output. The ability to combine these different types of knowledge is not constant but can be expected to vary systematically during particular phases of the evolution of the lifecycle. For example, we can expect the ability to combine different types of knowledge to fall as a completely new type of

knowledge emerges as a discontinuity and to rise again as the new type of knowledge starts maturing.

To yield the knowledge coherence index, a number of steps are required. The following calculations show how to obtain the index at whatever level of analysis i . First of all, we need to calculate the parameter τ , i.e. *technological relatedness*, by deriving the relatedness matrix as follows (Nesta, 2008). Let the technological universe consist of k patent applications. Let $P_{jk} = 1$ if the patent k is assigned the technology j [$j = 1, \dots, n$], and 0 otherwise. The total number of patents assigned to technology j is $O_j = \sum_k P_{jk}$. Similarly, the total number of patents assigned to technology m is $O_m = \sum_k P_{mk}$. Since two technologies may occur within the same patent, $O_j \cap O_m \neq \emptyset$, the number of observed co-occurrences of technologies j and m is $J_{jm} = \sum_k P_{jk} P_{mk}$. By applying this relationship to all possible pairs, we yield a square matrix Ω ($n \times n$) whose generic cell is the observed number of co-occurrences:

$$\Omega = \begin{bmatrix} J_{11} & & J_{j1} & & J_{n1} \\ \vdots & \ddots & & & \vdots \\ J_{1m} & & J_{jm} & & J_{nm} \\ \vdots & & & \ddots & \vdots \\ J_{1n} & \dots & J_{jn} & \dots & J_{nn} \end{bmatrix} \quad (14)$$

We can assume that the number x_{jm} of patents assigned to both technologies j and m is a hypergeometric random variable of mean and variance:

$$\mu_{jm} = E(X_{jm} = x) = \frac{O_j O_m}{K} \quad (15)$$

$$\sigma_{jm}^2 = \mu_{jm} \left(\frac{K - O_j}{K} \right) \left(\frac{K - O_m}{K - 1} \right) \quad (16)$$

If the observed number of co-occurrences J_{jm} is larger than the expected number of random co-occurrences μ_{jm} , then the two technologies are closely related: the fact the two technologies occur together in the number of patents x_{jm} is not casual. The measure of relatedness hence is given by the difference between the observed number and the expected number of co-occurrences, weighted by their standard deviation:

$$\tau_{jm} = \frac{J_{jm} - \mu_{jm}}{\sigma_{jm}} \quad (17)$$

It is worth noting that this measure of relatedness has lower and upper bounds: $\tau_{jm} \in]-\infty; +\infty[$. Moreover, the index shows a distribution similar to a t-student, so that if $\tau_{jm} \in]-1.96; +1.96[$, one can safely accept the null hypothesis of non-relatedness of the two technologies j and m . The technological relatedness matrix Ω' may hence be thought of as a weighting scheme to evaluate the technological portfolio of regions.

Once the parameter τ has been calculated for each pair of technologies, we can proceed to derive the weighted average relatedness WAR_j of technology j with respect to all other technologies present within the relevant aggregate (say sector, firm or region). Following Teece et al. (1994), WAR_j is defined as the degree to which technology j is related to all other technologies $j \neq m$ in the aggregate, weighted by patent count P_{mit} :

$$WAR_{jit} = \frac{\sum_{m \neq j} \tau_{jm} P_{mit}}{\sum_{m \neq j} P_{mit}} \quad (18)$$

Finally, the coherence of knowledge base within the aggregate i (be it a firm, a sector or a region) is defined as the weighted average of the WAR_{it} measure:

$$R_{it} = \sum_{j \neq m} WAR_{jit} \times \frac{P_{jit}}{\sum_j P_{jit}} \quad (19)$$

It is worth stressing that the index implemented by analysing co-occurrences of technological classes within patent applications measures the degree to which the services rendered by the co-occurring technologies are complementary to one another (see Nesta and Saviotti, 2005, 2006 and Krafft, Quatraro and Saviotti, 2014). The relatedness measure τ_{jm} indicates that the utilization of technology j implies that of technology m in order to perform specific functions that are not reducible to their independent use. This makes the coherence index appropriate for the purposes of this study.

In addition to coherence, we also investigate the relationship between the terms of shift-share and **cognitive distance** (CD) (Nooteboom, 2000), which expresses the dissimilarities between different types of knowledge. This measure is of fundamental importance when distinguishing the effect of the emergence of a discontinuity from that of the subsequent period of normal or incremental development.

A useful index of distance can be derived from the measure of *technological proximity*. It was originally proposed by Jaffe (1986 and 1989), who investigated the proximity of firms' technological portfolios. Subsequently, Breschi et al. (2003) adapted the

index in order to measure the proximity, or relatedness, between two technologies. The idea is that each firm is characterized by a vector V of the k technologies that occur in its patents. Knowledge similarity can first be calculated for a pair of technologies l and j as the angular separation or un-centred correlation of the vectors V_{lk} and V_{jk} . The similarity of technologies l and j can then be defined as follows:

$$S_{lj} = \frac{\sum_{k=1}^n V_{lk} V_{jk}}{\sqrt{\sum_{k=1}^n V_{lk}^2} \sqrt{\sum_{k=1}^n V_{jk}^2}} \quad (20)$$

The idea underlying the calculation of this index is that two technologies j and l are similar to the extent that they co-occur with a third technology k . The cognitive distance between j and l is the complement of their index of the similarity:

$$d_{lj} = 1 - S_{lj} \quad (21)$$

Once the index is calculated for all possible pairs, it needs to be aggregated at the industry level to obtain a synthetic index of technological distance. This can be done in two steps. First of all, we can compute the weighted average distance of technology l , i.e. the average distance of l from all other technologies.

$$WAD_{lt} = \frac{\sum_{j \neq l} d_{lj} P_{jt}}{\sum_{j \neq l} P_{jt}} \quad (22)$$

where P_j is the number of patents in which the technology j is observed. Now the average cognitive distance at time t is obtained as follows:

$$CAD_{it} = \frac{\sum_l WAD_{lt} P_{li}}{\sum_l P_{li}} \quad (23)$$

Table 1 provides the synthetic definition of the variables which we have described so far.

>>> INSERT TABLE 1 ABOUT HERE <<<

4 Methodology and Data

4.1 Methodology

The main focus of this paper is on the observation of the co-evolutionary dynamics between knowledge and economic structure. In the previous section we proposed a synthetic representation of change in economic structure by introducing shift share analysis.

In view of the complex and endogenous nature of the relationships between the properties of knowledge and those of economic structure, we apply a VAR model.

The regression of interest is the following:

$$w_{i,t} = c + \beta_z w_{i,t-z} + \varepsilon_{i,t} \quad (24)$$

where w_{it} is an $m \times 1$ vector of random variables for region i at time t , β is an $m \times [m \times z]$ matrix of slope coefficients that are to be estimated. In our particular case $m=9$ and corresponds to the vector $[\mu(i,t), \pi(i,t), \alpha(i,t), \text{growth of knowledge capital (i,t), coherence growth (i,t), growth of cognitive distance (i,t), variety growth (i,t), related variety growth (i,t), unrelated variety growth (i,t)}]$. ε is an $m \times 1$ vector of disturbances.

In line with previous studies, the measure of growth rates is based on the difference in the logarithms of the respective variables. Let $X_i(t)$ represent the absolute value of the variable in region i at time t . Define the normalized (log) value of the variable as:

$$x_i(t) = \log(X_i(t)) - \frac{1}{N} \sum_{i=1}^N \log(X_i(t)) \quad (25)$$

Where N is the number of regions. In what follows, growth rates are defined as the first difference of normalized (log) values according to:

$$g_i(t) = x_i(t) - x_i(t-1) \quad (26)$$

In such a way, common macroeconomic shocks are already controlled because the growth rate distribution was normalized to zero for each variable in each region in each year.

Following a growing body of literature (Coad, 2010; Buerger, Broekel and Coad, 2012; Colombelli, Krafft and Quatraro, 2014a), Equation (18) is estimated via ‘reduced form’ VARs which do not impose any a priori causal structure on the relationships between the variables and are therefore suitable for the purposes of this analysis. These reduced-form VARs effectively correspond to a series of m individual ordinary least squares (OLS).

However, previous studies have emphasized how the empirical distribution of the growth rates is closer to a Laplacian than to a Gaussian distribution (Bottazzi et al. 2007; Bottazzi and Secchi 2003; Castaldi and Dosi 2009). Such evidence suggests that standard regression estimators, such as ordinary least squares (OLS), assuming Gaussian residuals may

perform poorly if applied to these empirical frameworks. To cope with this, a viable and increasingly used alternative consists of implementing least absolute deviation (LAD) techniques that are based on the minimization of the absolute deviation from the median rather than the squares of the deviation from the mean.

It must be noted that we do not include any individual dummies in the analysis. Even though unobserved heterogeneity can have important effects on the estimation results, the inclusion of individual dummies along with lagged variables may engender some biases for fixed-effect estimation of dynamic panel-data models, a problem known as Nickell-bias. Some alternative approaches relate to the use of instrumental variable (IV) or GMM estimators (Blundell and Bond, 1998). The main problem with this lies in the difficulty in finding good instruments which is particularly hard when dealing with growth rates. When instruments are weak, IV estimation of panel VAR thus leads to imprecise estimates. Binder et al. (2005), on the other hand, propose a panel VAR model including firm-specific effects which is, however, based on the assumption of normally distributed errors which is not the case for the growth rates of the variables used in our regressions.

Since we are dealing with growth rates rather than of levels, we can assume that any region-specific component has been largely removed. Moreover, we follow the wide body of literature on analysis of firms' growth rates which states that the non-Gaussian nature of growth rate residuals is a far more important econometric problem deserving careful attention even in regional level analyses (Buerger, Broekel and Coad, 2012).

4.2 The Data

In order to implement the methodology outlined in the previous section, we gather together two datasets. The shift-share analysis has been conducted by using the branch accounts of NUTS II European regions⁶ provided by the Eurostat within the European System of Integrated Economic Accounts. As is well known, these data have only been available

⁶ We acknowledge that the use of administrative regions to investigate the effects of knowledge creation represents only an approximation of the local dynamics underpinning such a process. Indeed, administrative borders are arbitrary and therefore might not be representative of the spontaneous emergence of local interactions. It would be much better to investigate these dynamics by focusing on local systems of innovation. However, it is impossible to find out data at such a level of aggregation. Moreover, the identification of local systems involves the choice of indicators and threshold values according to which one can decide whether to unbundle local institutions or not. This choice is in turn arbitrary and therefore would not solve the problem, but would only reproduce the issue at a different level. Thus, we think that despite the unavoidable approximation, our analysis may provide useful information on the dynamics under scrutiny.



since 1995, the year in which the Eurostat implemented a standardized procedure to collect data from European countries in order to build a coherent and homogeneous dataset. As a result, we were able to calculate the μ , the π , and the α components for a subset of European regions on a time span ranging from 1995 to 2007. The properties of knowledge structure, i.e. coherence, cognitive distance and variety (based on the information entropy index) have instead been calculated by using patent information contained in the OECD REGPAT database which covers patent data that have been linked to regions utilizing the addresses of the applicants and inventors. Analysis was conducted by adopting inventor-based regionalization⁷, and by using 4-digit technology codes.

We merged the two sets of indicators on the basis of the NUTS II regional code and the year. We ended up with an unbalanced panel of 227 regions observed on average over 8 years. The descriptive statistics for the whole sample are reported in Table 2 whereas Figure 1 shows the distributional properties of the variables under scrutiny, providing empirical support for their non-Gaussian distribution. In particular, all the variables appear to follow a Laplace-like distribution which makes our empirical strategy outlined in the previous section the best approach to the analysis.

>>> INSERT Figure 1 AND Table 2 ABOUT HERE <<<

The elaboration of a regional breakdown of descriptive statistics turns out to be very complicated when dealing with a sample of 227 regions. For this reason, we decided to show the cross-regional distribution of average values by implementing a map for each of the variables under consideration. The maps reported in Figure 3 and Figure 4 are based on absolute (log) values of the properties of the knowledge structure.

In Figure 2, we report the cross-regional distribution of the three components contributing to labour productivity growth. Let us recall that μ is the contribution of the changing mix of regional industries and is positive if regions tend to specialize in high-productivity activities, π is the interaction between productivity growth and the change in the industry mix and is positive to the extent that regions specialize in fast growing sectors

⁷ The assignment of patents to regions on the basis of inventors' addresses is the most widespread practice in the literature (see, for example, Maurseth and Verspagen, 2002; Henderson et al., 2005; Breschi and Lissoni, 2009, Paci and Usai, 2009, to quote a few). A viable alternative may rest on the use of applicants' addresses, above all when the assessment of knowledge impact on growth is at stake (see Antonelli, Krafft and Quatraro, 2010). However, when analysis is conducted at a local level of aggregation, and the geography of collective processes of knowledge creation is emphasized, the choice of inventors' addresses remains the best one.



whereas α is the contribution of within-sector productivity growth weighted by the sector share on total employment.

>>> INSERT Figure 2 ABOUT HERE <<<

It is interesting to note that for most of the sampled regions, the effect of change in the industry mix is positive, suggesting that structural change plays an important role in the process of economic growth. Most European regions therefore tend to specialize in high-productivity sectors, with the only exception being some Greek regions and the British Midlands. The process is more pronounced in Italy and in central-eastern Europe than in Spain and France. The second diagram shows that the interaction term is positive again in most Italian regions, Spain, France and Germany, while evidence is more mixed in other regions. Italy, France, Spain and Germany in the observed period are subject to changes favouring the increasing share of fast-growing sectors. Finally, within-sector productivity growth seems to matter the most for Northern regions, such as Finland, Sweden and Denmark, and to a somewhat lesser extent for some Eastern and Greek regions.

In Figure 3, we report the cross regional distribution of knowledge capital, coherence and cognitive distance (log values). The top diagram reports the figures concerning the knowledge capital. We can see how knowledge capital is higher in central European regions and in northern regions whereas it is lower in the periphery of the continent. A look at the coherence index reveals that on average search behaviours are more like organized search than random screening, while cognitive distance is on average very low in most of the European regions, suggesting that exploration is conducted across the safe boundaries of established knowledge competences. We observe both low values of coherence and cognitive distance only for a few scattered regions in France, Spain and Finland suggesting search strategies characterized by exploration behaviours conducted within well defined boundaries of the knowledge space.

>>> INSERT Figure 3 ABOUT HERE <<<

Figure 4 shows cross distribution of the variety index, articulated in unrelated and related knowledge variety. The top diagram indicates that on average European regions are characterized by a high degree of variety, with the only exception being some peripheral regions in Portugal and Greece. When we look at the distinction between related and unrelated variety, we can see that the distribution looks very similar to that of total variety. By



also observing the ranges assigned to each classes, we can also state that on average related knowledge variety is higher than unrelated variety.

>>> INSERT Figure 4 ABOUT HERE <<<

Table 2 shows the correlation across the variables used in the empirical analysis. It can be noted that only in a few cases correlation coefficients appear to be significantly correlated, but in any case the coefficients are pretty low.

>>> INSERT TABLE 2 ABOUT HERE <<<

Since the relationships under scrutiny are investigated at the regional level, the spatial dependence of the relevant variables can engender some biases in the estimation results. In order to check the extent to which spatial dependence can be a problem, we calculated the Moran's I index for all variables by relying on a row standardized contiguity matrix (Moran, 1950). As is well known, the Moran's I index is used to calculate the spatial autocorrelation across OLS residuals. Average growth rates were used because they reflect the fundamental relations between the regions in contrast to fluctuating yearly growth rates. The results are shown in Table 4, which suggests that only in two cases the statistics are significant and hence the null hypothesis of no spatial correlation ought to be rejected. However, the correlation coefficients are fairly low and this should not bias the subsequent estimations in any serious way. This is even more so considering that LAD estimation techniques are preferred over standard OLS.

>>> INSERT TABLE 4 ABOUT HERE <<<

In the following section we will implement the estimation of equation (18), which is based on the normalized growth rates of such variables.

5 Econometric results

The results of the econometric estimations are reported in Tables 5 to 8. Tables 5 and 6 present the results of exploratory regressions conducted by implementing the OLS estimator. Table 5 shows the estimations in which the shift share components are dependent variables and Table 6 shows the estimations in which the properties of the regional knowledge base are dependent variables.



With regard to the observed autocorrelation, it is impressive to note that none of the variables under scrutiny shows any degree of persistence. On the contrary, coefficients are negative and significant across all three lags considered, suggesting erratic growth dynamics for all the variables. Such results on knowledge-related variables are consistent with the findings of Buerger, Broekl and Coad (2012) who ascribe this kind of evidence to intrinsic uncertainty and volatility characterizing innovation.

>>> INSERT Table 5 ABOUT HERE <<<

We will now analyze in more detail the lead-lag relationship between the change in knowledge and economic structure. Let us start with Table 5. As far as the first lag is concerned, knowledge coherence (COH) and knowledge capital (KCAP) show a positive and significant coefficient on α which is consistent with the literature linking knowledge structure and productivity growth (Nesta, 2008; Quatraro, 2010). The α component stands for the contribution stemming from within-sector productivity growth which is positively affected by the growth of knowledge coherence and that of knowledge capital. Knowledge coherence is indeed likely to increase as a technological trajectory gets established and innovation capabilities begin to spread to regional innovating agents. Technological learning enhances the capacity to identify and retain profitable combinations of technologies in the generation of new knowledge. At the aggregate level, the knowledge base appears to be therefore characterized by a high degree of integration as a result of a cumulative process. Agents are now able to focus on a well defined area of the knowledge space and to engage in successful innovation efforts that are more likely to exert relevant effects on economic performances. The knowledge variety (KV, RKV and UKV) indexes as well as CD do not seem to affect the economic structure significantly.

When we move to the second lag, we see that knowledge coherence (KOH) affects the cross-term π positively and significantly, i.e. faster growth of coherence is associated with a faster increase in faster growing sectors. Once again, the ability to screen the knowledge space in an organized and systematic way allows for improved economic performance. Knowledge variety, on the other hand, negatively affects the cross-term. These negative coefficients can be interpreted in the light of the mixed nature of the cross-term π . We would



expect increasing technological variety to affect the change in the industry mix positively. On the contrary, higher growth rates of technological variety are typical of instability phases of the technological trajectory which are more likely to be associated with productivity slowdowns.

Finally, the third lag presents an interesting positive and significant coefficient of KV and UKV on the reallocation term (μ). This is consistent with the idea that increasing technological variety is likely to be associated with changes in the industry mix. The displacement of the workforce from mature activities to new emerging activities is indeed related to an increase in the scope of technological activities, focusing in particular on the combination of pieces of knowledge that belong to different technological domains.⁸

The effect of KV, RKV and UKV on the cross-term π is again negative, signalling the prevalence of the negative effects on productivity dynamics. Finally, it is also worth noting the positive and significant coefficient of knowledge coherence (COH) and cognitive distance (CD) as far as effects on within-sector productivity (α) are concerned. The former is in line with the evidence observed when looking at the first lag. Coherence is likely to increase as an effect of organized search activities typical of exploitation dynamics. These in turn are likely to engender significant productivity gains. The coupling of positive effects of coherence and cognitive distance suggest that productivity gains are associated with a combination of complementary but not similar pieces of knowledge. This is in line with the literature in the evolutionary economic geography field which suggests that too much similarity is detrimental to regional economic development (Boschma and Wenting, 2007; Colombelli and Quatraro, 2013)⁹.

>>> INSERT TABLE 6 ABOUT HERE <<<

Table 6 presents the results of the OLS estimation concerning the determinants of growth rates of the properties of regional knowledge bases.

⁸ The displacement effect on employment of course is unlikely to manifest itself in the shortest run. This may be the reason why this effect is grasped only when looking at the third lag. However, one would expect this evidence to become more robust in the subsequent lags. Unfortunately, the available data do not allow us to go further back in time.

⁹ See Colombelli, Krafft and Quatraro (2014a) and Colombelli and Quatraro (2014) for similar evidence at the firm level.



The evidence regarding the first lag shows that knowledge coherence (COH) appears to be positively affected by π , suggesting that the increasing share of fast growing sectors is associated with increasing integration of regional technological activities. It is also interesting to note that within-sector productivity (α) negatively affects the growth of cognitive distance. This evidence is in line with previous work (Colombelli, Krafft and Quatraro, 2014a) according to which higher performances are likely to create the economic conditions to stimulate patterns of knowledge creation based on the search across domains that are not so familiar and yet not too far from the established technological competences. The positive and significant effect of α on KV and RKV provides further support for such an interpretation.

The coefficients for the second lag of the explanatory variables reveal that the reallocation term (μ) does not yield any significant effect on the knowledge characteristics. The same applies to the cross-term (π). In line with the evidence for the first lag, within-sector productivity (α) has a negative effect on cognitive distance whereas it positively affects KV.

As far as the third lag is concerned, we can see that no significant effects of the shift-share terms on the properties of the regional knowledge base can be devised. The only exception is the negative effect of the reallocation term (μ) on knowledge variety. It would therefore seem that increasing variety can foster the changing allocation of labour across sectors, but that this in turn is likely to be followed by a reduction in variety.

The evidence discussed so far suggests that the effects of the properties of the regional knowledge base on the terms of shift-share decomposition are broader than the effects that the latter yield on the former. However, as already discussed in Section 4, OLS estimations are likely to be biased due to the distributional properties of the variables at stake. For this reason, we estimated the reduced form VAR by applying the LAD estimator. The results are reported in Tables 7 and 8.

>>> INSERT TABLE 7 ABOUT HERE <<<

Table 7 presents the evidence concerning the determinants of the shift share components. In columns (1a)-(1c), the dependent variable is the reallocation term (μ). Over the three lags, the autoregressive coefficient is negative and significant suggesting erratic movement of this term. With regard to the effects of the properties of the regional knowledge base, only KV shows a positive and significant coefficient at the third lag, which was already

present in the previous OLS estimation. This is in line with expectations. The enlargement of the scope for recombination activities signals the opening up of new trajectories for exploitation which attract economic agents. In the aggregate this is like to be associated with displacement of the workforce from old and mature sectors to new ones. Columns (2a)-(2c), on the other hand, show the results for the cross-term (π). Here too, the autoregressive terms over the three lags are negative and significant. As far as the knowledge-related variables are concerned, none of the coefficients is significant in the first lag. When we move to the following lags, KV turns out to be negative and significant both at the second and the third lag whereas RKV and UKV show negative and significant coefficients only on the third lag. As in the previous estimations, these results can be interpreted as an outcome of the dominance of the productivity gain component of the cross-term. Finally, in columns (3a)-(3c), the dependent variable is within-sector productivity (α). Here too, the autoregressive terms are negative and significant over the three implemented lags. Out of the knowledge-related variables, only CD shows a significant (and positive) coefficient on the third lag. This is in line with the previous estimation and suggests that in order to observe productivity gains, the regional knowledge base ought to be the outcome of the recombination of dissimilar pieces of knowledge.

>>> INSERT TABLE 8 ABOUT HERE <<<

In Table 8, we report the results of the reduced-form VAR estimations in which the properties of the knowledge base are dependent variables. In columns (4a)-(4b), the dependent variable is COH. In the first lag, no significant effects of the knowledge structure are observed. In the second lag, the cross-term (π) shows a positive and significant coefficient. This can be interpreted as an outcome of the dominance of the effect of the reallocation component. After displacement of the workforce from one sector to another, as an outcome of increasing variety which signals exploration strategies, technological activities are oriented towards increasing integration of the knowledge base in the new fields. This interpretation is also supported by the negative and significant coefficient of within-sector productivity (α). A technology-lifecycle interpretation is in order. When increasing, growth rates are indeed the outcome of exploitation dynamics that allow for the selection of profitable trajectories. After some time, high growth rates are followed by decreasing coherence due to the need to identify new and unexploited technological opportunities which allows for further productivity gains in the future.



In columns (5a)-(5c) the dependent variable is KCAP. In this case, the only shift-share variable showing a statistically significant coefficient is the cross-term and this applies to all of the three investigated lags. The sign of the coefficients is always negative. This could be interpreted as an outcome of the dominance of the reallocation component of the cross-term. A change in the sector allocation of the workforce in one region is indeed a sign of maturity of established activities and a search for new avenues of development. In this phase, the technological competences associated with old sectors may have exhausted the technological opportunities leading to a reduction of knowledge production which is not yet counterbalanced by identification of new profitable trajectories. The evidence on CD reported in columns (6a)-(6c), is consistent with the regional technology lifecycle interpretation articulated so far. Indeed, we can observe a negative and significant coefficient of within-sector productivity (α) in the first lag and a negative and significant coefficient of the cross-term (π) in the third lag. According to the evolutionary economic geography approach, high levels of CD are likely to positively affect regional competitiveness. When productivity gains are achieved, most technological opportunities have been exploited and the scope for recombination in the knowledge space gets smaller and smaller. As an outcome, the average degree of similarity amongst combinable technologies increases. This could also explain the negative sign of the cross-term which combines both the reallocation and the productivity components. Insofar as increasing reallocation is associated with the exhaustion of technological opportunities in incumbent sectors, the narrowing of the set of profitable combinations leads to the increasing likelihood that technologies with a high degree of similarity will be combined. Columns (6), (7) and (8) report the estimations for KV, RKV and UKV respectively. As far as KV is concerned, we can observe that the within-productivity term (α) shows a positive and significant coefficient in the second lag, while the cross-term (π) shows a negative and significant coefficient in the third lag. With regard to RKV, the only significant shift-share variable is the reallocation term which is featured by a negative and significant coefficient in the second lag. The estimation on UKV shows that only the cross-term is characterized by a significant and positive coefficient in the first lag. The evidence concerning the variety variables provides further support for our interpretation. Productivity growth is likely to be followed by an increase in knowledge variety due to the need to look for new profitable trajectories. Consistently, displacement of the workforce from incumbent to new sectors is typical of exploration phases. When this happens, we will probably observe a



reduction in knowledge variety in the following periods as a result of the gradual selection of successful combinations.

6 Conclusions

In this paper we have conducted an exploratory analysis of the co-evolutionary patterns of knowledge and economic structure. The concept of knowledge structure has been in particular elaborated so as to qualify the regional knowledge base according to its average degree of similarity, complementarity and variety. Drawing upon a theoretical framework that stresses the dynamic nature of the interactions between these two components as well as the endogenous character of the change process, we decided to implement an empirical framework based on the recombinant knowledge approach in order to characterize the knowledge structure. We coupled this methodological approach with the shift-share technique which allows the effects of the change in economic structure to be grasped in a synthetic way and in particular, we focused on the changing allocation of the labour force across sectors.

The empirical analysis, given the dynamic effects feeding back from economic and knowledge structure and vice versa, was conducted by implementing a set of ‘reduced-form’ VAR estimations which allowed us to investigate the lead-lag relationships between the two systems without imposing any aprioristic causal structure. The results of the analysis are encouraging and call for further research in this direction, showing a clear interactive pattern between the two structures. Changes in knowledge structure that signal the patterns of knowledge creation based on organized search are likely to engender increasing within-sector productivity growth, while knowledge generation activities characterised by random screening across the knowledge space are likely to be followed by changes in the allocation of the labour force across sectors. Moreover, the increasing share of faster growing sectors appears to stimulate the establishment of knowledge creation patterns dominated by organized search within the comfortable fences of established competences. Moreover, econometric analysis also allowed us to appreciate some interesting dynamics such as the one relating



variety and μ , according to which increasing variety fosters the changing allocation of labour across sectors, but which in turn is likely to be followed by a reduction in variety¹⁰.

These results represent an interesting starting point for investigating and shedding new light on issues traditionally addressed in regional economics and economic geography such as the importance of the effects of the surrounding environment on firms' economic performances or the possible effects of the structure of regional (or local) knowledge bases on firms' location choices, which can be enriched by fully appreciating the role of the properties of local knowledge bases. These issues are also important to the policy realm, too often focused on the design of measures targeted to restoring regional competitiveness by providing funding to incumbent activities. Although measures targeted to restore productivity in incumbent sectors are less complicated, these are more likely to yield temporary outcomes. On the contrary, the results of our analyses suggest that policymakers should promote the entry of new activities grounded on new and profitable technological trajectories. Measures aiming at reshaping the structure of economic and technological activities at the local level are likely to yield enduring effects since they are based on structural interventions.

In this direction, the scope for demand-driven policy instruments at the local level is enriched. Indeed, new activities cannot emerge out of the blue. The incentives to the local creation of new technology based activities should be therefore grounded on the accurate analysis of both the comparative advantages developed over time in a specific area and of the relative position of such technologies in the technological landscape (Colombelli, Krafft and Quatraro, 2014b; Asheim and Coenen, 2005). Stimulating local agents to jump to new activities far away from their cumulated competencies can be inefficient and unsuccessful.

Strategic demand pull emerges here as an ingredient to the strategic management of places, the goal of which should be the promotion of knowledge-based entrepreneurship as a vehicle for the employment growth and global competitiveness at the local level (Audretsch, 2003). Demand-driven innovation policies should therefore be particularly cautious in the identification of the key sectors to be promoted, above all when implemented at the European level, as this latter task cannot be performed without a careful screening of the patterns of local technological specialization in the different areas upon which the policy instrument

¹⁰ An important limitation of this exercise concerns the lag structure used in the analysis. Although most of the existing studies using the reduced VAR approach do not go beyond the third lag (Coad and Rao, 2010; Buerger, Broekl and Coad, 2012), it would be useful to investigate longer lag structures. The available data on regional economic accounts unfortunately do not allow analysis to be extended in this direction.



impinges. These policy measures should be rather customized, so as to ensure effective implementation and the reduction of duplication of efforts and waste of resources.



7 References

- Acs, Z.J., Anselin, L. and Varga, A., 2002. Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, 31, 1069-1085.
- Antonelli, C., 2008. *Localized Technological Change: Towards the Economics of Complexity*. London: Routledge.
- Antonelli, C., Krafft, J. and Quatraro, F., 2010. Recombinant knowledge and growth: The case of ICTs. *Structural Change and Economic Dynamics* 21, 50-69.
- Antonelli, C. and Quatraro, F., 2013. Localized Technological Change and Efficiency Wages: the Evidence across European Regions. *Regional Studies* 47, 1469-1483.
- Asheim, B. T., & Coenen, L., 2005. Knowledge bases and regional innovation systems: Comparing Nordic clusters. *Research policy* 34, 1173-1190.
- Audretsch, D. B., 2003. Entrepreneurship policy and the strategic management of places. In D. M. Hart (Ed.) *The emergence of entrepreneurship policy: Governance, start-ups, and growth in the US knowledge economy*, 20-38. New York: Cambridge University Press.
- Binder, M., C. Hsiao and C. H. Pesaran, 2005. Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory* 21, 795-837.
- Blundell R.W and Bond S.R., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115-143.
- Boschma, R. 2005. Proximity and innovation: A critical assessment. *Regional Studies* 39: 61-74.
- Boschma, R. and Frenken, K., 2006. Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography* 6, 273-302.
- Boschma, R. and Frenken, K., 2011. Technological relatedness and regional branching. In: Bathelt, H., Feldman M.P. and Kogler, D.F. (ed.) *Dynamic geographies of knowledge creation and innovation*. London and New York, Routledge.
- Boschma, R. and Wenting, R., 2007. The spatial evolution of the British automobile industry: Does location matter? *Industrial and Corporate Change*, 16, 213-238.
- Bottazzi, G., Cefis, E., Dosi, G. and Secchi, A., 2007. Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries. *Small Business Economics* 29, 137-159.
- Bottazzi, G., Secchi, A., 2003. Common properties and sectoral specificities in the dynamics of U.S. manufacturing companies. *Review of Industrial Organization* 23, 217-232.
- Breschi, S. and Lissoni, F., 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography* 94, 439-468.
- Breschi, S., Lissoni, F., and Malerba, F., 2003. Knowledge relatedness in firm technological diversification, *Research Policy* 32: 69-97.



- Buerger, M., Broekel, T. and Coad, A., 2012. Regional dynamics of innovation: Investigating the co-evolution of patents, research and development (R&D), and employment. *Regional Studies* 46, 566-582.
- Burns A. F., 1934. *Production trends in the United States since 1870*. NBER, Boston.
- Castaldi, C. and Dosi, G., 2009. The patterns of output growth of firms and countries: Scale invariances and scale specificities. *Empirical Economics* 37, 475-495.
- Coad, A., 2010. Exploring the processes of firm growth: evidence from a vector auto-regression. *Industrial and Corporate Change* 19, 1677-1703.
- Coad, A. and Rao, R., 2010. Firm growth and R&D expenditure. *Economics of Innovation and New Technology* 19, 127-145.
- Colombelli, A., Krafft, J. and Quatraro, F., 2014a. High Growth Firms and Technological Knowledge: Do gazelles follow exploration or exploitation strategies?, *Industrial and Corporate Change* 23, 261-291.
- Colombelli, A., Krafft, J. and Quatraro, F., 2014b. The emergence of new technology-based sectors at the regional level: A proximity-based analysis of nanotechnologies. *Research Policy*, forthcoming.
- Colombelli, A. and Quatraro, F. 2013. New Firm Formation and the properties of local knowledge bases: Evidence from Italian NUTS 3 regions, Working Papers hal-00858989, HAL.
- Colombelli, A. and Quatraro, F. 2014. The persistence of firms' knowledge base: A quantile approach to Italian data, *Economics of Innovation and New Technology*, forthcoming.
- Creamer, D., 1942. *Shifts of Manufacturing Industries, in Industrial Location and National Resources*. Washington, D. C.: U. S. National Resources Planning Board.
- Dosi, G., 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy* 11, 147-162
- Esteban, J.M., 1972. Shift-share analysis revisited. *Regional and Urban Economics* 2, 249-61.
- Esteban, J.M., 2000. Regional convergence in Europe and the industry mix: a shift-share analysis. *Regional Science and Urban Economics* 30, 353-364.
- Fagerberg, J., 2000. Technological Progress, Structural Change and Productivity Growth: A Comparative Study. *Structural Change and Economic Dynamics* 11, 393-411.
- Fisher, A.G.B, 1939. Production, primary, secondary and tertiary. *The Economic Record* 15, 24-38.
- Fleming, L., 2001. Recombinant uncertainty in technological search. *Management Science* 47, 117-132.
- Fleming, L., Sorenson, O. 2001. Technology as a complex adaptive system: Evidence from patent data. *Research Policy* 30: 1019-1039.
- Frenken, K. 2004. Entropy and information theory. In: Hanusch, H, and Pyka, A. (ed.), *The Elgar Companion to Neo-Schumpeterian Economics*. Cheltenham, Edward Elgar, pp. 544-555.



- Frenken, K., Nuvolari, A., 2004. Entropy Statistics as a Framework to Analyse Technological Evolution. In: J. Foster and W. Hözl (ed.) *Applied evolutionary economics and complex systems*. Cheltenham, U.K. and Northampton, Mass.: Edward Elgar.
- Frenken, K., von Oort, F., Verburg, T., 2007. Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies* 41: 685-97.
- Griliches, Z., 1990. Patent statistics as economic indicators: A survey. *Journal of Economic Literature* 28, 1661-1707.
- Hall, B.H., Griliches Z. and Hausman J.A., 1986. Patents and R and D: Is there a lag?. *International Economic Review* 27, 265-283.
- Henderson, R. M., Jaffe, A. and Trajtenberg. M., 2005. Patent citations and the geography of knowledge spillovers: A reassessment: Comment. *American Economic Review* 95, 416-464.
- Houston, G.B., 1967. The Shift and Share Analysis of Regional Growth: A Critique. *Southern Economic Journal* 33, 577-581.
- Jaffe, A.B. 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review* 76, 984-1001.
- Jaffe, A., 1989. Real Effects of Academic Research. *American Economic Review* 79: 957-70.
- Krafft, J., Quatraro, F. and Saviotti, P.P., 2011. The knowledge-base evolution in biotechnology: a social network analysis. *Economics of Innovation and New Technology* 20, 445-475.
- Krafft, J., Quatraro, F. and Saviotti, P.P., 2014. Evolution of the knowledge base in knowledge intensive sectors. *Industry and Innovation*, forthcoming.
- Kuznets S., 1930. *Secular Movements in Production and Prices*. Houghton Mifflin, Boston.
- Lawson C. and Lorenz, E., 1999. Collective learning, tacit knowledge and regional innovative capacity. *Regional Studies* 33, 305-317.
- Lever, W. F., 1999. Competitive cities in Europe. *Urban studies* 36, 1029-1044.
- Maurseth, P. B. and Verspagen, B., 2002. Knowledge spillovers in Europe: A patent citations analysis. *Scandinavian Journal of Economics* 104, 531-45.
- Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37, 17-23.
- Nelson, R.R. and Winter, S., 1982, *An Evolutionary Theory of Economic Change*, Cambridge, Harvard University Press.
- Nesta, L., 2008. Knowledge and productivity in the world's largest manufacturing corporations. *Journal of Economic Behavior and Organization* 67: 886-902.
- Nesta, L., and Saviotti, P.P., 2005. Coherence of the Knowledge Base and the Firm's Innovative Performance: Evidence from the U.S. Pharmaceutical Industry, *Journal of Industrial Economics* 53: 123-42.
- Nesta, L., and Saviotti, P.P., 2006. Firm Knowledge and Market Value in Biotechnology. *Industrial and Corporate Change* 15: 625-52.
- Nightingale, P., 1998. A cognitive model of innovation. *Research Policy* 27, 689-709.



- Nooteboom, B., 2000. *Learning and innovation in organizations and economies*. Oxford: Oxford University Press.
- Paci, R. and Usai, S., 2009. Knowledge flows across European regions. *The Annals of Regional Science* 43, 669-690.
- Pavitt, K., 1985. Patent statistics as indicators of innovative activities: Possibilities and problems. *Scientometrics* 7, 77-99.
- Perloff, H. S., Dunn, E. S., Lampard, E. E. And Muth, R. F., 1960. *Regions, Resources and Economic Growth*. Baltimore, Maryland: Johns Hopkins Press.
- Perroux F., 1955. Note sur la notion de 'pole de croissance'. *Économie Appliquée* 7, 307-320.
- Quatraro, F., 2009a. Innovation, structural change and productivity growth. Evidence from Italian regions, 1980-2003. *Cambridge Journal of Economics* 33, 1001-1022.
- Quatraro, F., 2009b. The diffusion of regional innovation capabilities: Evidence from Italian patent data. *Regional Studies*, 43, 1333-1348.
- Quatraro, F., 2010. Knowledge Coherence, Variety and Productivity Growth: Manufacturing Evidence from Italian Regions, *Research Policy*, 39, 1289-1302.
- Quatraro, F., 2012. *The economics of structural change in knowledge*. London and New York, Routledge.
- Romijn H. and Albu, M., 2002. Innovation, networking and proximity: Lessons from small high technology firms in the UK. *Regional Studies* 36, 81-86.
- Saviotti, P. P., 1988. Information, variety and entropy in technoeconomic development. *Research Policy* 17: 89-103.
- Saviotti, P.P., 2004. Considerations about the production and utilization of knowledge, *Journal of Institutional and Theoretical Economics* 160: 100-121.
- Saviotti, P.P., 2007. On the dynamics of generation and utilisation of knowledge: The local character of knowledge, *Structural Change and Economic Dynamics* 18: 387-408.
- Schumpeter, J. A., 1939. *Business Cycles. A Theoretical, Historical and Statistical Analysis of the Capitalist Process*. New York and London, McGraw Hill.
- Schumpeter. 1942. *Capitalism, Socialism and Democracy*, London, Unwin.
- Shannon, C.E., 1948. A mathematical theory of communication, *Bell System Technical Journal* 27: 379-423.
- Sorenson, O., Rivkin, J.W. and Fleming, L., 2006. Complexity, network and knowledge flows. *Research Policy* 35, 994-1017.
- Teece, D.J., Rumelt, R., Dosi G. and Winter, S., 1994. Understanding Corporate Coherence: Theory and Evidence. *Journal of Economic Behavior and Organisation* 23: 1-30.
- Theil, H., 1967. *Economics and Information Theory*. Amsterdam: North Holland.
- Tödting, F., & Trippel, M., 2005. One size fits all?: Towards a differentiated regional innovation policy approach. *Research policy* 34, 1203-1219.
- Usai, S., 2011. The Geography of Inventive Activity in OECD Regions. *Regional Studies*, 45, 711-731.



Weitzman, M. L., 1998. Recombinant growth. *Quarterly Journal of Economics* 113, 331-360.