RELATED VARIETY, UNRELATED VARIETY,
AND LOCAL *VS* GLOBAL SPILLOVERS.

Ernest Miguélez
GREThA, University of Bordeaux, & AQR-IREA & CReAM

Rosina Moreno
AQR-IREA, University of Barcelona

Abstract

The goal of this paper is twofold. First, within a regional knowledge production function (KPF) framework, we assess which diversified sectorial structure (related vs. unrelated variety) generates more knowledge spillovers which will ultimately enhance the innovation output. Different from previous studies, we distinguish between the local, inward-looking economy of regions and the sectors exposed to external interactions. This relates directly to the distinction between local and global knowledge spillovers. Second, we study in depth the relationships between the local knowledge economy and the sectors exposed to external interactions in generating new regional knowledge. In particular, we assess whether the more similar the internal and external knowledge sectors, the larger the innovation outputs, or else, different but related internal and external sectors are more prone to innovation creation. We use a sample of 274 NUTS2 European regions of 27 countries from 1999 to 2007.

Key words: variety, co-patenting, externalities, technological relatedness, knowledge production function
JEL: C8, J61, O31, O33, R0

1. INTRODUCTION

Motivating this paper is the belief that human skills and humans' social interactions drive the production and dissemination of ideas. This relevance lies in the principle that the combination and recombination of previously unconnected ideas lead to new knowledge production, subsequent technological innovations, and ensuing economic growth and well-being (Aghion and Howitt, 1992; Jones, 1995). Further, knowledge diffusion in the form of knowledge spillovers is central to this literature as a cause of endogenous growth (Romer, 1986, 1990) and geographic agglomeration of firms (Audretsch and Feldman, 1996, 2004; Jaffe et al., 1993).

A relevant issue in the knowledge externalities literature is whether firms located in agglomerations mainly learn from other local firms in the same industry or from other local firms in a range of other industries (Glaeser et al, 1992). The former dates back to Marshall's (1920) contributions on the benefits arising from spatial concentration, i.e., (1) markets for intermediate industries (input-output relationships within industries), (2) labour market pooling, as well as (3) positive externalities arising from other firms' knowledge stocks, or knowledge spillovers, that are 'in the air'. As Romer (1986, 1990) puts it several decades later, firms' output builds not only on their own inputs, but also on the stock of knowledge of the whole society.

The second form of externalities relate to Jane Jacobs' contributions on cities, externalities and innovation (Jacobs 1969; see also Glaeser et al., 1992). From her work we learn that a diversified economy would bring benefits to local firms because it would generate new ideas steaming from the cross-fertilization of ideas across different industries. In a regional context, this would be in line with the idea that regions with a more diverse stock of knowledge would have a greater potential for innovation and growth. However, since the paper by Frenken et al (2007), several authors have argued that the concept of diversification claimed by Jacobs need to be more deeply specified, by separating between diversification of related industries and diversification of unrelated industries – or, using the correct jargon, related versus unrelated variety. Regions hosting related industries, with different but connected knowledge bases, can more easily engage in recombinant innovation. On the contrary, the combination of previously unrelated industries or technologies is more difficult to succeed.

In parallel to this, a separate stream of literature in economic geography has discussed at length the importance of co-location and cities for the production of new ideas; at some point, co-located agents may start to combine and recombine local knowledge that eventually becomes redundant and less valuable. As a result, processes of lock-in may begin to occur (Arthur, 1989; Boschma, 2005; David, 1985). Conversely, firms looking for external sources of knowledge that lie beyond their own boundaries may find that the knowledge they require is available beyond the boundaries of the region (Bergman and Maier, 2009), which has challenged the traditional believes on the role of physical proximity. Hence, the interplay between a vibrant ‘local buzz’ and more intentional ‘global pipelines’ is important to ensure an optimal regional rate of innovation adoption and further knowledge creation (Bathelt et al., 2004)

We contribute to these different strands of literature in several ways. First, within a regional knowledge production function (KPF) framework, we assess which diversified sectorial structure (related vs. unrelated variety) generates more knowledge spillovers which will ultimately enhance the innovation output. Different from previous studies, we distinguish between the local, inward-looking economy of regions and the sectors exposed to external interactions. This relates directly to the distinction between local and global knowledge spillovers.

Second, we study in depth the relationships between the local knowledge economy and the sectors exposed to external interactions in generating new regional knowledge. In particular, we assess whether the more similar the internal and external knowledge sectors, the larger the innovation outputs, or else, different, but related internal and external sectors are more prone to innovation creation.

Our methodological approach builds upon the large literature analysing the impact of variety (related and unrelated) on economic outcomes (see section 2 for a throughout review), with few differences. First, given our interests, we compute diversity and variety indexes using a technological classification of economic activities, which turns out to be more meaningful for our purposes. This is in contrast with the majority of studies, which define the variables for variety in economic terms, that is, computing diversity indexes with data on either employment or imports (such as in Boschma and

Iammarino 2009; Tavassoli and Carbonara 2014). Information of the technological classification contained in patents from the European Patent Office (EPO) is used to build the different indexes. In addition, among the scarce papers investigating cross-regional linkages and related variety, trade data are used –either imports or exports between each pair of regions (Boschma and Iammarino, 2009). Instead, we use information on one specific channel of knowledge flow, namely research collaborations proxied through the co-patenting structure provided by patent documents. According to previous empirical literature, co-inventorship in patents are a good proxy of knowledge flows between regions that enhances innovation output (Miguélez and Moreno, 2013a, 2013b).

The outline of the paper is as follows: section 2 reviews the related literature. Section 3 offers the explanation of the empirical analysis and in section 4 the data are described. We give the main results in section 5 and finally we conclude in section 6.

2. RELATED LITERATURE

Much research in the field of economic geography and regional science deals with the impact of agglomeration economies on growth and put the emphasis on the question whether specialization in few sectors or diversity is more relevant in enhancing growth (Jacobs, 1969; Glaeser et al, 1992). In particular, the advantages of specialization have been argued using the notion of "localization externalities" and are, in particular, mainly due to the presence of specialized input suppliers, a local pool of specialized labour skills and specialized knowledge concerning the particular industry (Marshall, 1920). These external economies are available to all local firms within the same sector. On the other hand, diversification into a large variety of sectors (Jacobs' externalities) implies an improvement in the opportunities to interact, copy, modify and recombine ideas, practices and technologies across industries.

However, the concept of diversity presents a higher complexity, as first signaled by Frenken et al (2007). These authors pose the central question of whether it is related or unrelated diversification which is most relevant for growth. Unrelated variety measures the extent to which a region is diversified in very different types of activities; whereas related variety would be the best measure for Jacobs' externalities because it measures the variety within each of a class of technology/activity. According to their results, unrelated variety protects a region against external asymmetric shocks in demand and thus against rising unemployment, since the risk is spread over unrelated sectors. On the contrary, related variety in a sector enhances growth and employment, since it is beneficial for Jacobs' externalities in the form of knowledge spillovers.

These first results from Frenken et al (2007) in favour of the significance of related variety for regional growth in the Netherlands were confirmed by studies in other countries: Bishop and Gripaios (2010) for Great Britain, Boschma and Iammarino (2009), Quatraro (2010) for Italy, Hartog et al (2012) for Finland and Boschma et al (2012) for Spain. However, the role of unrelated variety is more controversial: whereas Bishop and Gripaios (2010) find that unrelated variety is affecting employment growth in a larger set of industries than related variety, Boschma et al (2012) and Hartog et al (2012) do not find any growth effect for unrelated variety. Frenken et al (2007) find that unrelated variety dampens unemployment growth since a wide range of unrelated industries in a given region spreads risks, given that when a shock occurs in a sector, it is scarcely likely to affect negatively other sectors and disturb the regional economy.

In this paper we aim at analysing the role of variety on regional innovative performances. Initially, one would expect that the variety of the knowledge stock within a region plays a role in the generation of knowledge spillovers as it is associated to the Schumpeter's notion of novelty by combination of previous ideas. In addition, it is also expected that such variety of knowledge can have a higher impact on innovation if it is a related variety rather than an unrelated variety. This is not a novelty in the literature since Tavassoli and Carbonara (2014) and Castaldi et al (2014) also analyse the role of related and unrelated variety on the regional innovation output, for the Swedish and the US cases, respectively. However, only the paper by Castaldi et al use the information contained in patent documents about the technological fields in order to construct the measures of regional variety in the knowledge stock of a region. We will do so for all the information concerning the NUTS2 regions in the European Union. We expect that knowledge originated in sectors different from those in which the region is specialized, but related, will enable effective connections since related technologies are more easily recombined. On the other hand, in the case of knowledge originated in sectors that are very different from the specialization of the region (unrelated variety), the knowledge base would not be able to easily absorb it so that little spillovers would be generated. However, this unrelated variety of the regional knowledge stock could imply connecting previously unrelated technologies which could support breakthrough innovation (Castaldi et al, 2014).

We also plan to complement the ideas of Jaffe (1989) and Feldman and Florida (1994) who consider that the relevant knowledge for many local firms is knowledge that spills over from other local R&D activities, external to the firm but still internal to the region. In contrast, we put to the forefront an important debate within the geography of innovation literature that has emerged recently, that is, the role of external knowledge linkages in the process of regional knowledge creation. Indeed, the widely accepted assumption that firms usually source their innovations from their immediate vicinity, might have limited our understanding of the ways in which knowledge flows across space and the way in which innovations are generated (Coe and Bunnell 2003). Thus, recent empirical works have extensively documented the influence of extra-local knowledge sources on firms' innovative performance and knowledge acquisition (Owen-Smith and Powell 2004; Gittelman 2007; Gertler and Levite 2005; Rosenkopf and Almedia 2003; Simonen and McCann 2008).

In addition, authors such as Boschma (2005), Camagni (1991) and Grabher (1993) highlight the increasing importance of firms' needs to network with extra-local knowledge pools to overcome potential situations of regional 'entropic death' or 'lock-in'. Otherwise, subsequent local interactions lead to the combination and recombination of the same pieces of knowledge, and firms would end up stuck in strong social structures that tend to resist social change (Boschma and Frenken 2010; Morrison et al., forthcoming) and prevent them from recognizing opportunities in new markets and technologies (Lambooy and Boschma 2001). On the contrary, **'**distant contexts can be a source of novel ideas and expert insights useful for innovation processes (…). Firms therefore develop global pipelines not only to exchange products or services, but also in order to benefit from outside knowledge inputs and growth impulse' (Maskell et al. 2006, p. 998).

As a consequence of the ideas above, in this paper we will also analyse the role of extra-regional linkages in the process of knowledge creation. Specifically, we will focus on one specific mechanism of knowledge flow or diffusion, namely, the establishing of patenting collaborations between inventors residing in different regions (across-region co-patenting). Networks of inventors may imply cross-pollination of previously unconnected ideas that will lead to better knowledge outputs (Katz and Martin 1997). Besides, individuals connected within a collaborative framework are more willing to learn from each other than is the case of isolated inventors. Moreover, collaborative research projects may achieve scale economies and may lessen research costs by reducing duplication of efforts and decreasing uncertainty among the participants in the network (Powell and Grodal 2005).

Research on collaborative research networks, and their impact on knowledge diffusion and innovation, has expanded greatly in recent years. This is particularly true in the case of networks of co-inventors, thanks to the availability of relevant data (that is, co-patent data). Part of this literature has been devoted to explaining the determinants of these collaborative patterns (Hoekman et al. 2009; Maggioni and Uberti, 2009), while another important strand has focused on networks as mechanisms for inter-regional R&D spillovers (Kroll, 2009; Ponds et al. 2010), and, in particular, networks as the means by which knowledge diffuses between individuals and across firms (Breschi and Lissoni, 2004, 2006; 2009; Gomes-Casseres et al., 2006; Singh, 2005).

As a consequence of the stream of ideas surveyed above, and thanks to the information on co-patenting across European regions, we plan to analyse if the variety of the knowledge stock of a regional economy has a different impact on the generation of innovation if this variety is given in knowledge which is produced exclusively by inventors within the region or if it is produced in interaction with inventors out of the region. We also plan to check if the extra-regional knowledge flows generated in research collaboration networks should be related to, but not the same as the knowledge base of a region in order to get the highest benefit. A priori, we expect that the degree of relatedness should be neither too small, to avoid lock-in, nor too large, to facilitate the absorption of such extra-regional knowledge. This analysis will allow us to offer empirical evidence on the role of the relatedness between sectors as a source of external economies by exploring the channels through which such spillovers may occur, namely co-patenting in our case.

3. EMPIRICAL ANALYSIS

3.1 Empirical model and measures

We use a KPF framework at the regional level. Our point of departure is the simplest specification of this model:

$$Y = f(RD, Z), \tag{1}$$

where Y is the innovative output of a given region, which depends on regional R&D expenditures (RD) as well as Z, a number of time-variant controls that account for specific features of the region i at time t. Among them, the measures of variety, as explained in the following subsections. Moreover, regional differences in size are accounted for by dividing the dependent and explanatory variables by total population. All in all, the following model is suggested:

$$\mathbf{ln} Ypc_{it} = \beta \cdot \mathbf{ln} RDpc_{it} + Z_{it} + \delta_i + \delta_t + \varepsilon_{it}, \tag{2}$$

where $ln Ypc_{it}$ is the log-transformation of the annual number of patent applications per million inhabitants, $ln RDpc_{it}$ is the log-transformation of R&D expenditures per capita, and Z are, again a number of focal variables – as explained below – and controls. In additional, δ_i and δ_t stand for, respectively, regional time-invariant fixed-effects and time fixed-effects. In order to consider deviations from the theory, a well-behaved error term is also introduced, ε_{it}.

A specialization index and a concentration index of industries constructed using patents from 30 IPC[1] technological sectors –OST subdivision- are also included, in order to control for the influence of specialization and concentration economies on innovation (Feldman and Audretsch 1999). To calculate the technological specialization index, we employ the following formula

---

[1] International Patent Classification

$$\mathrm{SpIn}_{it} = \frac{1}{2}\sum_{j}\left|\frac{\mathrm{PAT}_{ijt}}{\mathrm{PAT}_{it}} - \frac{\mathrm{PAT}_{Cjt}}{\mathrm{PAT}_{Ct}}\right|, \tag{3}$$

where PAT is the number of patents in each region i for each sector j, expressed as a difference for the whole sample of regions (C). The concentration index is built as follows:

$$ConIn_{it} = \sum_{jt}\left(PAT_{ijt} / PAT_{jt}\right)^{2}. \tag{4}$$

3.2 Related and unrelated variety

In the present paper, several entropy measures are employed to investigate empirically the quantitative relationship between related and unrelated variety and the levels of inventiveness. In line with previous papers, we measure related and unrelated variety with entropy measures at different levels of sectoral aggregation using the patenting profile of each region.

Following Frenken et al (2007), unrelated variety per region is proxied by the entropy of the two-digit distribution. Formally, if all three-digit sectors j fall under a two-digit sector S_g, where g=1,…, G, it is possible to derive the two digit shares, P_g, by summing the three-digit shares p_j

$$P_g = \sum_{j \in S_g} p_j \tag{5}$$

The unrelated variety measure is given by the entropy at the two-digit level:

$$UV = \sum_{g=1}^{F} P_g log_2\left(\frac{1}{P_g}\right) \tag{6}$$

and we consider that it measures the extent to which a region is diversified in very different types of activities.

Related variety is indicated by the weighted sum of the entropy at the three-digit level within each two-digit class:

$$RV = \sum_{g=1}^{G} P_g H_g \tag{7}$$

where:

$$H_g = \sum_{j \in S_g} \frac{p_j}{P_g} log_2 \left(\frac{1}{p_j / P_g} \right) \tag{8}$$

It measures the diversity of a region's portofolio at the most fine disaggregation. In this sense, it can be understood as a measure of Jacobs' externalities.

The indices of related and unrelated variety are not opposites. One region can have both a high related variety (diversified into many specific classes in each big category) and a high unrelated variety (diversified into unrelated technological categories). In fact, they tend to correlate positively (Frenken et al, 2007; Boschma et al, 2012), although it is not always the case.

Following with the empirical model sketched above, we include now the indices proxying for related and unrelated variety in the Z vector including controls that account for specific features of the region,

$$Z_{it} = g(RV_{it}, UV_{it}), \tag{9}$$

Inserting the related and unrelated variety indexes into the main equation yields to:

$$\mathbf{ln} Ypc_{it} = \beta \cdot \mathbf{ln} RDpc_{it-1} + \lambda_1 SpIn_{it-1} + \lambda_2 ConIn_{it-1} + \omega_1 RV_{it-1} + \omega_2 UV_{it-1} + \delta_i + \delta_t + \varepsilon_{it}, \tag{10}$$

We introduce the subscript t-1 to all the explanatory variables in order to make clear that they have been time lagged one period to lessen endogeneity concerns due to system feedbacks. Section 4 includes further details regarding the construction of all the variables used in the present analysis.

3.3 Relatedness in external interactions

Regions are not isolated entities that do not interact with the rest of the world; rather, an increasing number of studies have documented the fact that firms in regions source more and more their innovations in non-local knowledge interactions. Indeed, we consider that spatial networks formation is also likely to be conducive to knowledge diffusion, knowledge recombination and innovation. At the level of European regions, Ponds et al. (2010) and Maggioni and Uberti (2009) show the importance of cross-regional interactions to the process of regional innovation. However, little attention has been paid to the study of which kind of external interactions may be more beneficial. Indeed, the literature on agglomeration economies has considered the role of intersectoral linkages within the region, but little is done across regions. Merging both strands of research, we conjecture that new variety may enter a region thanks to the interactions with other regions. As such, we hypothesise that extra-regional knowledge flows should be related to some extent to the sectoral specialization of a region to have an impact, but should not be the same.

To determine the similarity between the external knowledge that enters a region through interactions and its regional specialization we compute a knowledge similarity index. It is computed as the sum of the products of the absolute sizes of the three-digit industry's patents ($PAT_3(j)$), as a proxy of the knowledge stock in a region, and extra-regional co-patents in it ($COPAT_3(j)$), proxying or the knowledge that enters from other regions:

$$KNOWSIM = log \sum_{j} PAT_3(j)\, COPAT_3(j) \tag{11}$$

This measure gets a maximum when the region is specialized in just one industry and the same one for patents and extra-regional co-patents. The lowest values are obtained when the more diverse the region is in patents and extra-regional co-patents and the less similar both profiles are. The idea is that when a region gets knowledge from other regions thanks to research collaboration interactions but such knowledge comes from the same industry the region is specialized in, the knowledge base of the economy will be able to absorb it but it will not add much to the existing knowledge.

The degree of relatedness between the knowledge base in the region and the knowledge that enters from other regions through co-patenting activities is measured through the indicator RELATEDNESS following the same idea as in Boschma and Iammarino (2009):

$$RELATEDNESS = \sum_{j} COPAT_3^M(j)\, PAT_3(j) \tag{12}$$

where $COPAT_3^M(j)$ is the co-patenting entropy in 3-digit technologies other than j, but within the same two-digit industry, and $PAT_3(j)$ is the relative size of the 3-digit patent technology j in the total regional patenting. The idea is that for each 3-digit patent technology in a region (e.g., technology 215), we measured the entropy of the co-patents from the other 3-digit technologies (e.g., technologies 211, 212, 213, 214 and 216) within the same 2-digit class (technology 21), excluding the same 3-digit co-patent industry (i.e., technology 215).

4. DATA

We use a sample of 274 NUTS2 European regions of 27 countries – EU-27 (except Cyprus and Malta) plus Norway and Switzerland, to estimate a regional KPF from 1999 to 2007).

Our dependent variable, innovation output, is measured by patent applications, a variable widely used in the literature to proxy innovation outcomes. As widely documented, this proxy presents serious caveats since not all inventions are patented, nor do they all have the same economic impact, as they are not all commercially exploitable (Griliches 1991). In spite of these shortcomings, patent data have proved useful for proxying inventiveness as they present minimal standards of novelty, originality and potential profits, and as such are a good proxy for economically profitable ideas (Bottazzi and Peri 2003). We retrieve patent data at the regional level from the OECD REGPAT database – July 2013 edition (Maraut et al., 2008).

As for the explanatory variables, R&D expenditures data (both private and public expenditures in regions) were collected from Eurostat and some National Statistical Offices, with some elaboration for regions in specific countries (Belgium, Switzerland, Greece, and the Netherlands). R&D expenditures, as well as the remaining covariates, are time-lagged one period in order to lessen endogeneity problems due to system feedbacks.

We use unit-record data retrieved from EPO patents – OECD REGPAT database, July 2013 edition – to construct the co-patenting variable on which some of the main measures in this paper are constructed. Table 1 provides summary statistics of the variables used in the present analysis.

[Insert Table 1 about here]

5. RESULTS

5.1 Results on the impact of variety within the region

We estimate an unbalanced panel FE model of 9 periods (from 1999 to 2007, both inclusive). Using longitudinal data and including FE in our regressions allow us to improve previous estimates of the KPF key parameters – to the extent that these FE account for a number of time-invariant unobservable characteristics of the regions that might bias the results if not included. Table 2 provides the two-way fixed effects estimates for the regional KPF model, including R&D per capita as well as a specialization and a concentration index, in addition to our focal variables, namely the variety indicators. Column (i) uses as dependent variable the logarithm of patents, fractional count, per million inhabitants of the region. In column (ii) the endogenous is the full count of the sum of patents in the region per capita, whereas column (iii) is the full count of patents weighted by the number of citations the patent has received in subsequent patent documents, in a time window of 3 years after the application year.. This is an attempt to proxy for the quality of the patents, understanding that the more citations a patent receives, the higher its value (Tranjtenberg, 1990).

[Insert Table 2 about here]

In all the cases, the Hausman test rejects the null hypothesis that individual effects are uncorrelated with the independent variables, so the fixed-effects model is preferred to the expense of the random-effects – results provided upon request. In general, the KPF holds in the European regional case for the period under consideration. The elasticity of patents with respect to R&D expenditures presents significant values (0.15-0.20), which is in line with the value obtained in the literature (Jaffe 1989; Bottazzi and Peri 2003).

With respect to the variety index for the European regional case, we also obtain that the variety in the knowledge stock of a region is indeed positively and significantly related to its innovation output. This is the case in the three columns. Therefore, it seems that a region with higher variety can profit from higher learning opportunities thanks to the idea of “novelty by combination” (Schumpeter, 1934). However, once we separate between related and unrelated variety, both parameters are positive although only the first one is significant. This points to the fact that the higher the number of technologically related sectors in a region, the higher across-sector knowledge externalities and likely the more learning opportunities for them (Frenken et al. 2007). Thus, the learning opportunities generated by variety of activities within the region are relevant when such activities are related so that knowledge externalities between such related sectors are generated. We can therefore conclude that the main hypothesis is confirmed: it is not just having a diversified economy what matters, but rather an economy that comprises related activities (related in terms of competences) that generate knowledge spillovers.

5.2 Results on the impact of cross-regions externalities and their composition

So far, the sources of knowledge externalities due to the variety of sectors/technologies are external to the firms but still internal to the region. Next, in Table 3 we assess to what extent the variety of the knowledge stock of a region has a different impact on the innovation output if this variety is given in the knowledge produced exclusively by inventors within the region (internal sector) or produced in interaction with inventors out of the region (external sector). As it can be observed, the main hypothesis of the role of related variety holds for the internal sector, whereas the results change for the external one: it is unrelated variety in the knowledge stock of the patents invented in collaboration with out-of-the-region inventors which impacts significantly positive whereas it is no longer the case for related variety.

[Insert Table 3 about here]

In any case, the role of extra-regional knowledge flows generated in research collaboration networks can be more directly analysed through the indices that explicitly consider to what extent the knowledge that flows from other regions is related to, but not the same, as the knowledge stock of the host region. Table 4 shows the results when the similarity and relatedness indices are included. As observed, similarity between the technological composition of the knowledge of the within-the-region patents and that of the cross-regional co-patents has no significant impact on the regions' innovative output. In other words, if the knowledge that enters a region thanks to collaboration agreements with inventors in other regions is from sectors in which it already patents, there is little room for absorbing new knowledge, with the subsequent non-significant impact on innovation output.

[Insert Table 4 about here]

We obtain that it is the degree of relatedness between the technological sectors of the patents in the region and the technological sectors of the knowledge flows that come from co-patenting with inventors in other regions what matters. The positive and significant parameter of the relatedness index suggests that the connections made with inventors outside the region have a positive and significant impact on its innovation output as far as the co-patenting profile (the knowledge that enters the region from the external world) and the knowledge stock of the region are related. This is probably the case because the possibilities to learn different competences are wider in such a case.

If we compute the indices of similarity and relatedness between the technological sectors in within-the-region patents and the technologies in the international collaborations (that is, discarding collaborations with inventors in other regions within the same country), the result for relatedness is maintained whereas the patent similarity is now negatively significant. Indeed, the similarity variable already presents some negative influences in previous literature, although in such cases (Boschma and Iammarino 2009; Tavassoli and Carbonara 2014) the authors compute similarity indexes between imports and exports, which is always measured as international trade linkages. This is therefore more related to the similarity index we are constructing now, in which only international collaborations are taken into account.

6. CONCLUSION

The value added of this paper is the study of how the relatedness between industries works as a source of external economies in the generation of new knowledge in the European Union by exploring the differences when the knowledge is generated within or beyond the frontiers of the region. Specifically, we assess whether the relatedness of certain networking activity across EU regions, measured through co-patenting, affect regional innovation, since it may bring new and related variety into the region.

REFERENCES

Aghion, P., Howitt, P., 1992. A Model of Growth Through Creative Destruction. Econometrica 60, 323. doi:10.2307/2951599

Arthur, W.B. (1989) Competing Technologies, Increasing Returns, and Lock-in by Historical Events, Economic Journal 99: 116-131.

Audretsch, D.B. and Feldman, M.P. (1996) R&D Spillovers and the geography of innovation and production, American Economic Review 86(3): 630-640.

Audretsch, D.B. and Feldman, M.P. (2004) Knowledge Spillovers and the Geography of Innovation, in V. Henderson and J. Thisse (Eds) Handbook of Urban and regional Economics, Volume 4.

Bathelt, H., Malmberg, A., Maskell, P., 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. Prog Hum Geogr 28, 31–56. doi:10.1191/0309132504ph469oa

Bergman, E.M. and Maier, G. (2009) Network central: regional positioning for innovative advantage, The Annals of regional Science 43(3): 615-644.

Bishop, P. and Gripaios, P. (2010) Spatial externalities, relatedness and sector employment growth in Great Britain, *Regional Studies* 44(4), 443–454.

Boschma, R. (2005) Proximity and Innovation: A Critical Assessment, Regional Studies 39(1): 61-74.

Boschma, R.A. and Iammarino, S. (2009) Related Variety, Trade Linkages and Regional Growth, *Economic Geography* 85(3): 289-311.

Boschma, R. and Frenken, K. (2010) The spatial evolution of innovation networks. A proximity perspective, in: R. Boschma and R. Martin (eds.) The Handbook of Evolutionary Economic Geography, Cheltenham: Edward Elgar, pp. 120-135.

Boschma, R., Minondo, A. and Navarro, M. (2012) Related variety and regional growth in Spain, *Papers in Regional Science* 91(2), 241–256.

Bottazzi, L. and Peri, G. (2003) Innovation and spillovers in regions: Evidence from European patent data, *European Economic Review* 47(4): 687-710.

Breschi, S. and Lissoni, F. (2004) Knowledge networks from patent data: methodological issues and research targets. In Moed H, Glänzel W, Schmoch U. (eds) Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems, pp. 613–643. Springer Verlag, Berlin.

Breschi, S. and Lissoni, F. (2006) Cross-firm inventors and social networks: localised knowledge spillovers revisited, *Annales d'Economie et de Statistique* 79-80: 189-209.

Breschi, S. and Lissoni, F. (2009) Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows, *Journal of Economic Geography* 9: 439-468.

Camagni, R (1991) Local 'milieu', uncertainty and innovation networks: towards a new dynamic theory of economic space, ch 7, pp 121-142 in Camagni, R (ed) Innovation Networks: Spatial Perspectives, Belhaven Press, London.

Castaldi, C., K. Frenken and B. Los (2014) Related Variety, Unrelated Variety and Technological Breakthroughs: an analysis of U.S. state-level patenting, Regional Studies, 41(5), 685-697.

Coe, N.M. and Bunnell, T. (2003) 'Spatializing' knowledge communities: towards a conceptualisation of transnational innovation networks, *Global Networks* 3(4): 437-456.

David, P.A. (1985) Clio and the Economics of QWERTY, American Economic Review 75(2): 332-337.

Feldman, M.P., Audretsch, D.B., 1999. Innovation in cities:: Science-based diversity, specialization and localized competition. European Economic Review 43, 409–429. doi:10.1016/S0014-2921(98)00047-6

Feldman, M. P. and R. Florida. (1994) The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation in the United States, Ann*als of the Association of American Geographers*, 84(2): 210-229.

Frenken, K., F. van Oort and T. Verburg (2007) Related Variety, Unrelated Variety and Regional Economic Growth, Regional Studies, 41(5), 685–697.

Gertler, M.S. and Levite, Y.M. (2005) Local nodes in global networks: the geography of knowledge flows in biotechnology innovation, *Industry and Innovation* 12(4): 487-507.

Gittelman, M. (2007) Does geography matter for science-based firms? Epistemic communities and the geography of research and patenting in biotechnology, *Organization Science* 18(4): 724-741.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A. and Shleifer, A. (1992) Growth in Cities, Journal of Political Economy 100(6): 1126-1152.

Gomes-Casseres, B., Hagedoorn, J., Jaffe, A.B., 2006. Do alliances promote knowledge flows? Journal of Financial Economics 80, 5–33. doi:10.1016/j.jfineco.2004.08.011

Grabher, Gernot. (1993). The Weakness of Strong Ties. The Lock-in of Regional Development in the Ruhr Area. In: The Embedded Firm, ed. G. Grabher. London: Routledge, pp. 255-277.

Griliches, Z. (1991) Patent Statistics as Economic Indicators: A Survey. NBER Working Papers No. 3301. National Bureau of Economic Research.

Hartog, M., Boschma, R. and Sotarauta, M. (2012) The impact of related variety on regional employment growth in Finland 1993-2006: High-tech versus medium/lowtech, *Industry and Innovation* 19 (6), 459-476.

Hoekman, J., Frenken, K., Tijssen, R.J.W., 2010. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. Research Policy 39, 662–673. doi:10.1016/j.respol.2010.01.012

Jacobs J (1969) The economy of cities. Random House, NewYork.

Jaffe, A.B., 1986. Technological Opportunity and Spillovers of R & D: Evidence from Firms’ Patents, Profits, and Market Value. The American Economic Review 76, 984–1001. doi:10.2307/1816464

Jaffe, A.B., 1989. Real Effects of Academic Research. The American Economic Review 79, 957–970. doi:10.2307/1831431

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. The Quarterly Journal of Economics 108, 577–598. doi:10.2307/2118401

Jones, C. (1995) R&D-Based Models of Economic Growth, Journal of Political Economy 103: 759-784.

Katz, J.S., Martin, B.R., 1997. What is research collaboration? Research Policy 26, 1–18. doi:10.1016/S0048-7333(96)00917-1

Kroll, H. (2009) Spillovers and proximity in perspective: A network approach to improving the operationalisation of proximity. Working Papers Firms and Regions No. R2/2009, Fraunhofer ISI.

Lambooy, J.G. and Boschma, R.A. (2001) Evolutionary economics and regional policy, *The Annals of Regional Science* 35(1): 113-133.

Maggioni, M., Uberti, T., 2009. Knowledge networks across Europe: which distance matters? The Annals of Regional Science 43, 691–720.

Maraut S, Dernis H, Webb C, Spiezia V, Guellec D (2008) The OECD REGPAT Database: A presentation STI Working Paper 2008/2.

Marshall, A. (1920) Principles of Economics. Londres, Macmillan.

Maskell, P., Bathelt, H. and Malmberg, A. (2006) Building global knowledge pipelines: The role of temporary clusters, *European Planning Studies* 14(8): 997-1013.

Miguélez, E., Moreno, R. (2013a). Skilled labour mobility, networks and knowledge creation in regions: a panel data approach. The Annals of Regional Science 51, 191–212. doi:10.1007/s00168-012-0526-0

Miguelez,E. and R. Moreno (2013b) Do labour mobility and technological collaborations foster geographical knowledge diffusion? The case of European regions, *Growth and Change*, 44(2), 324-357.

Morrison, A., Rabellotti, R. and Zirulia, F.L. (2011) When do global pipelines enhance knowledge diffusion in clusters?. Papers in Evolutionary Economic Geography (PEEG) 1105, Utrecht University.

Owen-Smith, J. and Powell, W.W. (2004) Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community, *Organization Science* 15(1): 5- 21.

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

Powell, S. and Grodal, W.W. (2005) Networks of Innovators. In: The Oxford Handbook of Innovation, Fagerberg J, Mowery DC, Nelson RR (eds) Oxford University Press, Oxford.

Quatraro, F. (2010) Knowledge coherence, variety and productivity growth: Manufacturing evidence from Italian regions, *Research Policy* 39, 1289–1302.

Romer, P.M. (1986) Increasing Returns and Long-run Growth, Journal of Political Economy 94(5): 1002-1037.

Romer, P.M. (1990) Endogenous Technological Change, Journal of Political Economy 98(5): 71-102.

Rosenkopf, L. and Almeida, P. (2003) Overcoming Local Search Through Allainces and Mobility, *Management Science* 49: 751-766.

Simonen, J. and McCann, P. (2008) Firm innovation: The influence of R&D cooperation and the geography of human capital inputs, *Journal of Urban Economics* 64(1): 146-154.

Singh, J. (2005) Collaborative Networks as Determinants of Knowledge Diffusion Patterns, *Management Science* 51: 56-70.

Tavassoli, S. and N. Carbonara (2014) The role of knowledge variety and intensity for regional innovation, *Small Business Economics*, 43(2), 493-509.

Trajtenberg, M. (1990) A penny for your quotes: patent citations and the value of innovations, *RAND Journal of Economics* 20, 172-187.

Acknowledgements

The authors acknowledge financial support from Ayudas Fundación BBVA. Rosina Moreno also acknowledges the Program ICREA Acadèmia.

Table 1. Summary statistics

Variable	Obs	Mean	Std.Dev	Min	Max
PATpc Fractional	2466	107.59	130.84	0	1017.78
PATpcFull	2466	155.24	186.24	0	1122.39
PATWeighted	2466	247.67	304.18	0	2112.23
Variety	2466	5.65	1.74	0	7.78
Related Variety	2466	1.71	0.82	0	3.20
Unrelated Variety	2466	1.91	0.44	0	2.31
PATSIM	2466	4357.40	20171.26	0	256505.50
RELATEDNESS	2466	0.13	0.10	0	0.79
RD	2466	0.38	0.42	0.001	2.88
SPIN	2466	0.37	0.18	0.123	1.00
CONIN	2466	0.11	0.12	0	1.00

Table 2. Related and unrelated variety

	(1) FE	(2) FE	(3) FE	(4) FE	(5) FE	(6) FE
Variety	0.155**	0.155**	0.199**			
	(0.0726)	(0.0733)	(0.0845)			
Related Variety				0.172***	0.182***	0.221***
				(0.0533)	(0.0552)	(0.0669)
Unrelated Variety				0.0359	0.0343	0.0549
				(0.0322)	(0.0334)	(0.0381)
ln(R&D per capita)	0.185**	0.205**	0.197*	0.183**	0.203**	0.194*
	(0.0864)	(0.0905)	(0.109)	(0.0860)	(0.0900)	(0.109)
Specialization Index	-0.119**	-0.127**	-0.111	-0.127**	-0.133**	-0.118
	(0.0558)	(0.0610)	(0.0763)	(0.0512)	(0.0586)	(0.0741)
Concentration Index	0.0438*	0.0451	0.0318	0.0420*	0.0429	0.0302
	(0.0244)	(0.0291)	(0.0367)	(0.0235)	(0.0282)	(0.0359)
Constant	3.636***	3.971***	4.389***	3.631***	3.966***	4.385***
	(0.0250)	(0.0271)	(0.0327)	(0.0256)	(0.0274)	(0.0329)
Observations	2,466	2,466	2,466	2,466	2,466	2,466
R-squared	0.169	0.191	0.121	0.172	0.195	0.126
Number of regcode	274	274	274	274	274	274
Region FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Overall-R2	0.786	0.786	0.801	0.810	0.808	0.811
F-stat	21.67	24.61	14.54	20.82	23.56	14.49
F-prob	0	0	0	0	0	0

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dependent Variable (1): ln(patents p.c. (fractional counting) +1); D.V. (2): ln(patents p.c. (full counting) +1); D.V. (3): ln(cit. weight patents p.c. (full counting) +1) All explanatory variables are standardized

Table 3. Internal and external related and unrelated variety

VARIABLES	(1) FE	(2) FE	(3) FE
Related Variety External Sector	0.0178	0.0132	0.0256
	(0.0400)	(0.0461)	(0.0544)
Unrelated Variety External Sector	0.0707**	0.0646**	0.0676*
	(0.0310)	(0.0317)	(0.0391)
Related Variety Internal Sector	0.147***	0.153***	0.194***
	(0.0414)	(0.0418)	(0.0559)
Unrelated Variety Internal Sector	0.0243	0.0330	0.0541
	(0.0271)	(0.0320)	(0.0370)
ln(R&D per capita)	0.176**	0.197**	0.188*
	(0.0853)	(0.0895)	(0.108)
Specialization Index	-0.128***	-0.133**	-0.118
	(0.0488)	(0.0569)	(0.0728)
Concentration Index	0.0439*	0.0450	0.0314
	(0.0231)	(0.0276)	(0.0350)
Constant	3.634***	3.968***	4.383***
	(0.0261)	(0.0280)	(0.0341)
Observations	2,466	2,466	2,466
R-squared	0.177	0.199	0.132
Number of regcode	274	274	274
Region FE	yes	yes	yes
Time FE	yes	yes	yes
Overall-R2	0.802	0.792	0.795
F-stat	18.81	21.44	13.31
F-prob	0	0	0

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dependent Variable (1): ln(patents p.c. (fractional counting) +1); D.V. (2): ln(patents p.c. (full counting) +1); D.V. (3): ln(cit. weight patents p.c. (full counting) +1) All explanatory variables are standardized

Table 4. Similarity and relatedness between within-the-region knowledge and external knowledge

	(1)	(2)	(3)	(4)	(5)	(6)
	FE	FE	FE	FE	FE	FE
Variety	0.154**	0.154**	0.197**	0.151**	0.152**	0.195**
	(0.0726)	(0.0734)	(0.0845)	(0.0732)	(0.0739)	(0.0852)
Patent similarity	-0.0264	-0.0289	-0.0340			
	(0.0273)	(0.0278)	(0.0301)			
Relatedness	0.0643***	0.0713***	0.0650**			
	(0.0247)	(0.0229)	(0.0262)			
Patent similarity int'l sector				-0.0487**	-0.0504**	-0.0587**
				(0.0214)	(0.0215)	(0.0235)
Relatedness int'l sector				0.0216	0.0369***	0.0341**
				(0.0157)	(0.0134)	(0.0167)
ln(R&D per capita)	0.168*	0.187**	0.179	0.175**	0.193**	0.184*
	(0.0869)	(0.0912)	(0.110)	(0.0867)	(0.0906)	(0.109)
Specialization Index	-0.125**	-0.133**	-0.117	-0.118**	-0.127**	-0.111
	(0.0552)	(0.0603)	(0.0753)	(0.0559)	(0.0611)	(0.0764)
Concentration Index	0.0449*	0.0463	0.0328	0.0434*	0.0450	0.0315
	(0.0241)	(0.0286)	(0.0361)	(0.0244)	(0.0291)	(0.0367)
Constant	3.635***	3.970***	4.387***	3.630***	3.966***	4.382***
	(0.0256)	(0.0277)	(0.0335)	(0.0262)	(0.0283)	(0.0341)
Observations	2,466	2,466	2,466	2,466	2,466	2,466
R-squared	0.173	0.195	0.124	0.171	0.194	0.123
Number of regcode	274	274	274	274	274	274
Region FE	yes	yes	yes	yes	yes	yes
Time FE	yes	yes	yes	yes	yes	yes
Overall-R2	0.804	0.807	0.814	0.766	0.777	0.788
F-stat	19.69	21.90	12.87	19.17	21.89	12.74
F-prob	0	0	0	0	0	0

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Dependent Var. (1): ln(patents p.c. (fractional counting) +1); D.V. (2): ln(patents p.c. (full counting) +1); D.V. (3): ln(cit. weight patents p.c. (full counting) +1) All explanatory vars. are standardized