WP4/12 SEARCH WORKING PAPER

Interregional Knowledge Network Quality and Research Performance: Do Objective 1 and EU 12 Border Regions Follow Different Patterns than the Rest of Europe?

Tamás Sebestyén, Attila Varga

January 2013

SEARCH
SHARING KNOWLEDGE ASSETS: INTERREGIONALLY COHESIVE NEIGHBORHOODS

EUROPEAN COMMISSION
European Research Area
SEVENTH FRAMEWORK PROGRAMME
Funded under Socio-economic Sciences & Humanities

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2010-2.2-1) under grant agreement n° 266834

Interregional Knowledge Network Quality and Research Performance: Do Objective 1 and EU 12 Border Regions Follow Different Patterns than the Rest of Europe?

*Tamás Sebestyén^, Attila Varga**

^MTA-PTE Innovation and Growth Research Group and
Department of Economics and Regional Studies
Faculty of Business and Economics
University of Pécs

*Department of Economics and Regional Studies
MTA-PTE Innovation and Growth Research Group
Faculty of Business and Economics
University of Pécs
(Corresponding author)

Abstract
This paper estimates the impact of interregional knowledge flows on the productivity of research at the regional level with a particular attention given to Objective 1 and EU 12 border regions in this respect. The highlight of this latter aspect is related to the neighboring country focus of the SEARCH project. We apply the novel index of 'ego network quality' in order to measure the contribution of knowledge accessed from the interregional network to the production of new knowledge inside the region. Quality of interregional knowledge networks is related to the level of knowledge accumulated by the partners, the extent of collaboration among partners and the position of partners in the entire knowledge network. Ego network quality impacts on the productivity of research in scientific publications and patenting at the regional level are tested with co-patenting and EU Framework Program collaboration data for 189 European NUTS 2 regions. Though it is not possible to directly test the effects of co-patenting and FP network quality on EU neighboring countries we were able to get some estimates by testing the impacts of co-patenting and FP network quality on research productivity of NUTS 2 regions which possess the characteristics of EU neighboring countries in two respects: first, their GDP per capita is below the 75 % or EU average (Objective 1 regions) and second, they are located at the border of the old EU 12 territory (border regions).

Keywords: patents, scientific publications, knowledge networks, R&D productivity, regional knowledge production function, European regions, EU neighboring countries

JEL: O33, R11, R58

Interregional Knowledge Network Quality and Research Performance: Do Objective 1 and Border Regions Follow Different Patterns than the Rest of Europe?

1. Introduction

Scholarly attention towards the spatial dimension of knowledge flows has intensified in the past two decades in economics in general and in regional science and economic geography in particular. Since innovation is significantly related to existing knowledge, understanding the mechanisms of knowledge communication in space is essential for understanding the geography of economic growth. When knowledge flows (e.g., spillovers facilitated by informal relations, learning in research collaborations or knowledge transfers mediated by market transactions) are dominantly local, economic growth will most probably be highly uneven in space. Alternatively, globally flying knowledge may contribute to the spreading of economic growth (Fujita and Thisse 2002).

Most of the research in this field has focused on the role of spatial proximity in innovation. The early papers (Jaffe 1989, Jaffe, Trajtenberg and Henderson 1993, Anselin, Varga and Acs 1997) evidenced that knowledge flows between firms, private and public R&D labs are to a large extent localized geographically. Later on, several studies applying similar research methods in different countries supported this finding (Ghinamo 2012). It became clear soon, that besides pure knowledge spillovers, local labor markets (Breschi and Lissoni 2009), entrepreneurship (Zucker and Darby 1998) or formalized research collaborations (Miguélez and Moreno 2012) contribute importantly to the localized communication of knowledge. Though the proximity of industry-specific (specialized) knowledge could also be important, empirical findings suggest that the diversity of the local knowledge base is associated most frequently with regional growth (Glaeser, Kallar and Scheinkman 1991, Feldman and Audretsch 1999). Deeper analysis reveals that not a simple diversity of industries but their technological relatedness is what nurtures innovation (Frenken, van Oort and Verburg 2007).

Subsequent research clarifies that spatial proximity is neither a sufficient nor a necessary condition for knowledge flows to occur (Boschma 2005). Instead, other forms of proximities (such as cognitive, social, institutional or organizational proximities) are the crucial prerequisites of knowledge communication. Geographical proximity of actors in innovation may enhance the flow of knowledge but only if at least one of the other proximities are also in effect. Physically proximate location provides only the opportunities for frequent interactions. These interactions might become instrumental for the speedy flow of both tacit and codified knowledge as well as for the development of trust or the establishment of common codes of communication (Koschatzky 2000). Ponds, van Oort and Frenken (2009) show empirically that spatial proximity helps bridging institutional distances between business, academia and government. Thus, knowledge flows between different types of organizations tend to be localized. On the other hand, knowledge communication could also occur over larger physical distances among actors sharing similar institutional features.

Compared to the high intensity of research on localized knowledge interactions, scientific inquiry on the mechanisms and regional impacts of global knowledge flows is still a relatively less developed and novel phenomenon. Spatial econometric studies brought the first empirical evidence on the existence of interregional knowledge transfers and their positive role in regional innovation (Anselin, Varga and Acs 1997, Varga 1998). Recent efforts in the literature focus on studying various mechanisms of interregional knowledge flows ranging from professional labor mobility (Maier, Kurka and Trippl 2007, Schiller and Diez 2008, Miguélez, Moreno and Suriñach 2009), research collaboration (Maggioni and Uberti 2011) and co-inventorship (Breschi and Lenzi 2011, Broekel et al. 2010) to the operation of multinational companies (Cantwell and Iammarino 2003).

Network analysis (NA) is an especially promising tool for the study of interregional knowledge flows. Different measures of network structure such as network size, centrality of actors or density of interactions appear particularly powerful for understanding the geography of knowledge production. Much of the scientific inquiry in interregional knowledge interactions applying NA techniques grows out from the spatial econometrics tradition. Researchers realize that weights matrices routinely used in spatial econometrics to represent relations in space can also be applied to characterize relative positions in interregional knowledge networks (Maggioni and Uberti 2011). It is shown in this literature that the number of interregional partners in research or invention and their levels of knowledge (measured by e.g. R&D expenditures or publication stock) are indeed influential factors in regional knowledge production (Maggioni, Nosvelli and Uberti 2007, Hoekman, Frenken and van Oort 2009, Ponds, van Oort and Frenken 2009, 2010, Varga, Pontikakis and Chorafakis 2012).

Thus, the literature suggests that the larger the number of interregional partners is and the higher their knowledge levels are, the more effective a region becomes in knowledge generation. This is an important finding, but related to it at least two further issues arise. First, it can be argued that some substitutability might exist between network size and the level of knowledge in the network. A small network with a limited number of highly knowledgeable partners might be as valuable for a region as a large network with partners possessing diverse knowledge levels. While separate estimations of the impacts of the size and knowledge variables do not account for such a relationship in an econometric model, a comprehensive approach aggregating different network features into one combined measure could potentially be sensitive to such substitutability. Second, some additional characteristics of knowledge networks not yet considered in the geography literature (but accounted for in non-spatial studies) might also turn out to be influential in research productivity.

Building on Sebestyén and Varga (2013) in this paper we intend to make a step further in the state of art of the research addressing the relationship between interregional knowledge flows and regional knowledge creation. We structure the problem by directing attention to the *quality* of a particular interregional network for a particular region. This quality is reflected by the contribution of knowledge accessed in the network to the production of new knowledge inside the region. In order to make the problem empirically tractable we quantify certain features of interregional networks that could be instrumental for regional knowledge production additional to what have already been accounted for in the spatial literature (i.e., number of partners and their knowledge levels). Building on these measures and continuing the research line in Varga and Parag (2009) we then develop a novel comprehensive index to account for the quality of a region's global knowledge network. The higher this index value gets, the larger the amount of knowledge potentially accessed from the network will be. We also provide a test on the hypothesized positive relationship between interregional knowledge network quality and regional research productivity. In pursuit of this we investigate two interregional networks: EU Framework Program collaborations and European co-patenting, both considered at the level of EU NUTS2 regions. The two networks refer to two types of new knowledge generation processes: the production of new scientific knowledge measured by publications and the production of new technological knowledge measured by patent applications.

This paper estimates the impact of interregional knowledge flows on the productivity of research at the regional level with a particular attention given to Objective 1 and EU 12 border regions in this respect. The highlight of this latter aspect is related to the neighboring country focus of the SEARCH project. Though it is not possible to directly test the effects of co-patenting and FP network quality on EU neighboring countries we were able to get some estimates by testing the impacts of co-patenting and FP network quality on research productivity of NUTS 2 regions which possess the characteristics of EU neighboring countries in two respects: first, their GDP per capita is below the 75 % or EU average (Objective 1 regions) and second, they are located at the border of the old EU 12 territory (border regions).

Our paper is structured as follows. The second section introduces the concept of ego network quality, followed by the section describing the analytical model and the data used in the empirical analysis. The fourth section presents the results on how the quality of a region's interregional network affects the productivity of research in the creation of publications and patents at the regional level. In the fifth section we perform an analysis directly focusing on network quality impacts in Objective 1 and EU 12 neighboring regions. The last section summarizes our findings and highlights potential further research directions.

2. Ego network quality

Research on the interrelationships between the structure of actors' individual network and the performance of actors in knowledge production is a relatively recent phenomenon. Four features of individual (ego) networks are considered in this literature: 1) characteristics of immediate partners and the intensity of the actor's relations with them (number of partners, strength of ties, knowledge of partners); 2) intensity of interactions between partners; 3) diversity of knowledge accessed through the network (by being in contact with different fields of research); 4) the position of an agent in the entire network (in order to account for the impact of knowledge accessed beyond the ego network).

As shown in the literature, number of partners (Powell et al. 1999, Hopp et al. 2010, Van Der Deijl, Kelchtermans and Veugelers 2011), tie strength (Van Der Deijl, Kelchtermans and Veugelers 2011) and knowledge of partners (Maggioni, Nosvelli and Uberti 2006, Hoekman, Frenken and van Oort 2009, Ponds, van Oort and Frenken 2009, 2010, Varga, Pontikakis and Chorafakis 2012) are in a positive relationship with the productivity of research. However, the influence of the intensity of interactions is ambiguous. It positively affects agents' patenting productivity in Salmenkaita (2004) and Cross and Cummings (2004), the impact follows an inverted U-shape in Van Der Deijl, Kelchtermans and Veugelers (2011), while the influence on academic publishing is negative in Rumsey-Wairepo (2006) and Cainelli et al. (2010).[1] It is also found that both knowledge diversity of partners (Powell et al. 1999, Cainelli et al 2010, Van Der Deijl, Kelchtermans and Veugelers 2011) and a central position in the entire network (Powell et al. 1999, Cainelli et al. 2010, Hopp et al. 2010, Van Der Deijl, Kelchtermans and Veugelers 2011) positively affect performance in knowledge generation.

The theory of innovation emphasizes the role of interactions among different actors in innovation. These interactions follow a system and the characteristics of the system determines to a large extent the efficiency of new knowledge production (Lundvall 1992, Nelson 1993). An extensive survey-based empirical literature evidences that innovation is indeed a collective process where the knowledge and expertise of partners as well as the intensity of collaborations among them largely determines the production of new, economically useful knowledge (e.g., Diez 2002, Fischer and Varga 2002). Representing actors as nodes and their connections as ties, interactions of collaborating agents can be mapped as a network. On the basis of this representation the application of network analysis extends the frontiers of the study of knowledge interactions well beyond the possibilities of traditional innovation surveys.

In this paper we develop and apply the index of Ego Network Quality (ENQ) in empirically testing the impact of interregional knowledge networks on regional research productivity.[2] The larger the value ENQ gets, the higher the level of knowledge to be accessed from the network. Behind the concept of ENQ there are three intuitions directly influenced by the theory of innovation. The first intuition behind ENQ is that the level of knowledge in an agent's network is in a positive relationship with the agents' productivity in new knowledge generation. The second intuition is that collaboration among the partners in the agent's network is the source of further growth of knowledge available from the network. Following the third

[1] Referred to as intensity here, the question of cohesion in a network is also tackled from the social capital perspective. Here the debate is on wether cohesive, closed structures (Coleman, 1986) or 'structural holes' (Burt, 1992) provide a better background for performance. Although many of the results in this field show that a position in structural holes contribute to better performance in a diversity of fields (e.g. Hopp et al (2010), Kretschmer (2004), Donckels and Lambrecht (1997), Zaheer and Bell (2005), Powell et al. (1999), Tsai (2001), Burt et al. (2000), Burton et al (2010)), there is still evidence on the opposite (Salmenkaita, 2004), Cross and Cummings (2004). Rumsey-Wairepo (2006) argues that the two structural settings are complementaries rather than substitutes in explaining performance. In our context we also emphasize that different structural dimensions can be important for different networks. When information flows and power is important, structural holes indeed provide better position, however, as in our case, if knowledge production is in the focus, exclusion resulting from structural holes may be harmful and cohesiveness meaning better interaction may have positive contribution.

[2] Although the term ego network is usually used in a narrow sense in the network literature, meaning the subnetwork of a node (the ego) and its direct neighbors, we apply a broader perspective behind the notion, hence the term quality is used in the name. The main focus is on the level of knowledge attainable from the direct neighborhood, however, we take explicitly into account the fact that this quality does not exclusively depend on the network of direct neighbors but the network of their neighbors and so on. In a sense, if we would like to investigate the quality of the ego network, we must look behind direct partners because the quality of direct linkages crucially depend on the quality of the direct links of the neighbors and so on.

intuition we assume that partners in the ego network not only increase the amount of knowledge accessible, but also contribute to its diversity through building connections to different further groups not linked directly to the ego network.

ENQ integrates ego network features highlighted in the preceding paragraphs into one comprehensive measure. Based on the literature outlined above, we propose basically two dimensions for ENQ, which is then augmented with a related third aspect. The two dimensions are: Knowledge Potential and Local Connectivity. Knowledge Potential (KP) measures knowledge accumulated in the neighborhood and it is related to the number of partners and the knowledge of individual partners. Local Connectivity[3] (LC) is associated with strength of ties and the intensity of interactions among partners. The third aspect may be called Global Embeddeddness (GE) as it intends to capture the level and quality of knowledge developed in distant parts of the network (beyond immediate partners). However, this aspect is implemented by applying the concepts of KP and LC for consecutive 'circles' of indirect partners in the network.

The notation in the proceeding formulation is as follows. The network under question is formally represented by the adjacency matrix $\mathbf{A} = [a_{ij}]$, where the general element a_{ij} describes the connection between nodes i and j. These connections are considered as weights, normalized to the interval between 0 and 1. Thus a special case of this formulation is the binary network, where the elements of the adjacency matrix can be either zero or one. The adjacency matrix also defines the matrix of geodesic distances (lengths of shortest paths) between all pairs of nodes, which we denote by $\mathbf{R} = [r_{ij}]$.[4] In order to account for knowledge levels, we use $\mathbf{k} = [k_i]$ as the vector of knowledge at each specific node of the network.

A further note on terminology is required here. We frequently use the terms neighbor/neighborhood and distance in the paper. We are aware of the fact that these terms bear different meanings in network theory (neighbors and distances in the network) and regional science (neighbors and distances in space). As the focus of the paper is on the network approach, we use these terms in relation to neighbors and distances in the networks, especially in this section presenting ENQ. Later, when spatial relations are involved in the empirical analysis, we explicitly state if these terms are used in a spatial context.

Given the conceptual model presented above, we can formalize ENQ as follows:

$$ENQ^i = \sum_{d=1}^{M-1} W_d LC_d^i KP_d^i \qquad (1)$$

In this formula superscript i refers to the node for which ENQ is calculated and subscript d stands for distances measured in the network (geodesic distance). M is the size of the network, i.e. the number of nodes. W_d is a weighting factor used for

[3] Note that connectivity is used here in a broader sense than in graph theory. In graph theory connectivity refers to the number of vertices the removal of which disconnects the graph. In our case, this term refers to a similar concept but with a less strict definition. By connectivity we simply mean the extent of ties connecting a given group of vertices.

[4] In this paper we use a non-weighted algorythm for the calculation of geodesic distances, i.e. the distance of two nodes is regareded as the number of ties connecting them, irrespective of the weights associated to these ties.

discounting values at different d distances from node i,[5] whereas KP_d^i and LC_d^i are the respective Knowledge Potential and Local Connectivity measures evaluated in the network at distance d from node i.

Therefore, our proposed measure for ENQ is a distance-weighted sum of Local Connectivity-weighted Knowledge Potentials prevailing at different distances in the network. Moreover, the contribution of indirect partners at a given distance to the knowledge accessible by node i via its direct partners is calculated as Knowledge Potential weighted by Local Connectivity of partners at that distance. Applying a distinction between immediate and indirect partners in the network, we can reformulate ENQ as follows:

$$ENQ^i = W_1 LC_1^i KP_1^i + \sum_{d=2}^{M-1} W_d LC_d^i KP_d^i = LC_1^i KP_1^i + GE^i \qquad (2)$$

where we applied $W_1 = 1$ and the concept that everything beyond the immediate neighborhood can be labeled as Global Embeddedness (GE). In what follows, we give a detailed description of the two basic notions, Knowledge Potential and Local Connectivity.

Knowledge Potential

The concept of KP relates to the amount of knowledge an agent's partners possess. Using the notation presented before, the concept of KP can be formulated in the following way:

$$KP_d^i = \sum_{j:r_{ij}=d} k_j \qquad (3)$$

where d denotes geodesic distances measured in the network. The Knowledge Potential of node i can thus be calculated for different d distances (measured from node i), and for all these distances it is the sum of knowledge possessed by direct and indirect neighbors at these distances. As it was outlined before, this Knowledge Potential is going to be weighted by the Local Connectivity of direct and indirect neighbors. The concept of LC is presented in what follows.

Local Connectivity

As it was highlighted before, we assume that not only the knowledge levels of partners are of positive value to the node under question but also the cooperation between neighbors. More specifically we assume that each crosscutting tie has a positive value, depending on the weight the tie has. Local Connectivity is therefore the simple sum of the tie weights present in a given neighborhood, normalized by the size of these neighborhoods. The concept can be formulated as follows:

$$LC_d^i = \frac{1}{N_d^i}\left(\sum_{j:r_{ij}=d-1}\sum_{l:r_{il}=d} a_{jl} + \frac{\sum_{j:r_{ij}=d}\sum_{l:r_{il}=d} a_{jl}}{2}\right) \qquad (4)$$

The expression in the parenthesis consists of two parts. The first term counts the (weighted) ties between nodes at distance $d-1$ and d from node i. This reflects the

[5] A weighting factor is defined to be unity for $d = 1$ and descending towards zero as d increases. There is no unique best choice with regards the decay function. In this paper we employ a hyperbolic weighting, i.e. the decay function is $W_d = 1/d$.

intensity at which two 'circles' of neighbors are linked together. The second term counts the (weighted) number of ties among nodes at distance d from node i.[6] Thus, this is the intensity at which the (possibly indirect) neighbors at distance d are linked together. For the direct neighborhood of node i this expression sums the links among direct partners and the links connecting node i to its neighbors[7].

To better capture the specific meaning of the expression in equation (4) recall that we employ LC as a weighting factor to KP. Assume for example that node i has N_1^i direct partners and the links connecting it to these partners are of strength 1 ($a_{ij} = 1$ for all j in the direct neighborhood). If these partners have no connections among each other, then the second term in the parenthesis is zero, and the first term is N_1^i (because all connections have weight 1). Thus, LC is unity, which reflect the intuition that the knowledge levels of partners are fully absorbed. If the connections linking node i to its partners were less strong ($a_{ij} \leq 1$ for all j in the direct neighborhood) then LC would be lower than one, contributing to a lower weighting factor. This reflects the fact that in this case partners' knowledge is not fully accessible. Assume now, that the partners establish some links among each other. In this case the second term in the parenthesis starts to increase, and the weighting factor (LC) increases too. This reflects our previous concept, namely that a higher level of collaboration among partners contributes to the knowledge attainable from a network. To sum up, LC is a weighting factor, which describes how well connected the node is to its neighbors and how well these neighbors are connected to each other. However, the weighting is done according to a reference point: the weight is taken to be unity for the special case if links to the partners are of unit strength but no cooperation among partners is present.[8]

---

[6] Division by two is required because matrix $\mathbf{A}$ is symmetric, and thus we can avoid duplications in the counting. This division is not required in the first term because the definition there counts only links from distance $d-1$ to distance d and not vice versa.

[7] It is worth devoting a word to the inclusion of distance-crossing ties (the first term in the expression). Our intuition behind the concept of Local Connectivity is that collaboration among partners enhances knowledge sharing and this leads to a better environment for knowledge creation. If one looks at the direct network neighborhood, the links connecting the node in question and its neighbors becomes relevant in the general case of weighted ties: the amount of knowledge learned from the immediate partners depends on the intensity of interactions with those partners. However, if we look at more distant (indirect) neighborhoods (circles), it seems still important how well connected the nodes in these neighborhoods are, but the role of links bridging these neighborhoods together might not be that clear at first sight. We argue here that in our concept there is no significant difference between these links and those linking nodes at a given distance. The question at hand is that how well connected the neighborhood of a given node is, or in other words, that how dense the tissue of the network around the node is. We are going to attach less weigh to this connectivity the farther away it is from the node, but the main point is that better connectivity among nodes is of higher value, and this connectivity is not necessarily restricted to connectivity among nodes at a specific distance.

[8] A note is necessary here on non-direct neighborhoods, i.e. when $d > 1$. In these cases the normalization bears a different meaning from that in the direct neighborhood. In these indirect neighborhoods it is true that nodes at distance d must be connected with nodes at distance $d-1$ with at least as many links as many nodes there are at distance d, i.e. N_d^i. In this case, if all these links connecting nodes at distance $d-1$ and d are of unit strength, the first term in the parenthesis of equation (4) will be at least unity. However, it still holds that the weighting factor LC is unity in the special case if nodes at distance d are linked to nodes at distance $d-1$ through connections of unit strength *and* with the minimum number of connections required. It is also still valid that interconnections in the neighborhood at distance d increase the weighting factor and weaker connections between the different neighborhoods decrease the weighting factor. The only difference is

Ego Network Quality

The concepts of Knowledge Potential and Local Connectivity are linked together by weighting the Knowledge Potential with Local Connectivity. If we are looking only at the direct neighborhood of node i, we can write:

$$Q_1^i = KP_1^i LC_1^i \qquad (5)$$

This expression measures the average knowledge level of direct neighbors, weighted by the interaction among these neighbors.[9] However, as noted earlier, the level of knowledge attained from direct neighbors is enhanced by the level of knowledge these neighbors attain from their individual networks. Therefore, we augment our quality measure with the same connectivity-weighted average knowledge level of further indirect neighbors, using the distance weights as introduced before. According to these, the index of Ego Network Quality is defined as follows, which comes back to our starting definition in equation (1):

$$ENQ^i = \sum_d W_d Q_d^i = \sum_d W_d KP_d^i LC_d^i \qquad (6)$$

3. Empirical model and data

Our starting point is the knowledge production function initially specified by Romer (1990) and parameterised by Jones (1995). In the interpretation of the parameters we follow Varga (2006).

$$dA_i / dt = \delta \, H_{Ai}{}^{\lambda} A_i{}^{\varphi} \qquad (7)$$

where dA/dt is the temporal change in technological knowledge, H_A refers to research inputs (e.g. number of researchers or research expenditures), A is the total stock of already existing scientific and technological knowledge (knowledge codified into publications, patents etc.) and i stands for the spatial unit. Thus technological change is associated with contemporary R&D efforts and previously accumulated knowledge. The same number of researchers can have a varying impact on technological change depending on the stock of already existing knowledge. Two parameters are particularly important for this paper. The size of φ reflects the impact of the transfer of codified knowledge. Regarding parameter λ the larger its size is, the stronger the impact the same number of researchers will play in technological change. Its value thus reflects the productivity of research in region i. We assume that the size of λ is positively related to the quality of interregional knowledge networks measured by ENQ.

In order to test empirically the hypothesised relationships we follow Varga (2000) and Varga, Pontikakis and Chorafakis (2012) and use the following econometric specifications. Using subscripts *i* and *N* to denote individual regions and nations (in our case EU member states) respectively, the empirical counterpart of the Romer (KPF) is specified as:

that in these cases there is an extensive margin: the number of connections between the neighborhoods can also increase, and this increases the value of LC resulting in a higher weighting factor.

[9] Note, that the weight for $d = 1$ is unity by definition.

$$\log K_i = a_0 + a_1 \log RD_i + a_2 \log KSTOCK_N + Z_i + \varepsilon_i \quad (8)$$

where K stands for new scientific-technological knowledge, RD is expenditure on research and development, KSTOCK represents already existing scientific/technological knowledge at the national level and Z stands for additional regional control variables. We use national patent stock as a proxy for codified technological knowledge reachable with unlimited spatial accessibility within the country.

Equation 10 relates research productivity measured by $a_{1,i}$ the parameter of the research variable in Equation 9 to interregional network quality.

$$a_{1,i} = \beta_0 + \beta_1 \log ENQ_i \quad (9)$$

Substituting equation 9 to equation 8 results in the following equation to be estimated:

$$\begin{aligned} Log(K_{i,\,t}) = \alpha_0 + \beta_0 Log(RD_{i,\,t\text{-}k}) + \beta_1 Log(ENQ_i) {*} Log(RD_{i,\,t\text{-}k}) \\ + \alpha_2 Log(KSTCK_{N,\,t\text{-}k}) + Z_i + \varepsilon_i. \end{aligned} \quad (10)$$

The empirical analysis in this paper is based on a sample of 189 European regions (a mix of NUTS2 and NUTS1 regions) for which information was complete enough for our purposes. We use a cross-sectional database, though a two-year time lag is employed. The time period under consideration is determined by the duration of the 5th Framework Program of the EU, spanning the years 1998-2002. According to this, dependent variables are given for 2002, the independent variables are given for 2000 and the network variables are based on network connections observed throughout the period between 1998 and 2002.

With respect to knowledge flows in the production of new scientific results (resulting from Pasteur- type knowledge generation) on the one hand and new technological knowledge (a result of Edison-type knowledge production) on the other, we use different proxies for each, leading to different estimated equations and variables. Dependent variables are patenting activity and publication activity at the regional level as proxied by patent applications to the EPO (PAT_i) and scientific publications in ISI journals (PUB_i) respectively. We do not make any distinction between technological and scientific fields, so patent and publication counts are aggregated measures in this respect. Although using patents as a proxy for technological innovation is far from a perfect solution, there are several reasons why it still remains one of the most widely used and accepted measures (see e.g. Griliches 1990, for a comprehensive study on the issue, or Acs, Anselin and Varga 2002, for an analysis on the links between patent and other innovation counts at the level of regions). Publications are a somewhat stronger proxy for scientific knowledge and also used widely in innovation studies (van Raan 2004). Although the publishing motivation in academic science lends more reliability to publication counts, there are still biases stemming from journal coverage and distortions of evaluating mechanisms. Overall, although such biases may be relevant in an inter-regional comparison, given our central question here there are no strong reasons to think that it could affect European tendencies.

Table 1. Variable description

Variable Name	Description	Source
PAT_i	Number of patent applications to the European Patents office (EPO) by region of inventor (fractional counts)	Eurostat NewCronos database
PUB_i	Number of publications in scientific journals in the Thomson ISI database (search criteria: article, letter, review)	RFK database (data processed by CWTS, Leiden University)
GRD_i	Gross regional expenditures on R&D, in millions of Purchasing Power Standard (PPS) Euros, 1995 prices	Eurostat NewCronos database
$FPKP_i$ / $PATKPW_i$	Knowledge Potential – the directly available knowledge from a region's partners. FKP is calculated for the binary FP network with accumulated R&D stock as knowledge levels. PATKPW is calculated for the weighted patent network with accumulated patent stocks as knowledge levels.	Authors' elaboration on FP5 administrative database, DG RTD, Dir A and OECD REGPAT database
$FPLD_i$ / $PATLDW_i$	Local Density – the average number of links in a region's neighborhood. FPLD is calculated for the binary FP network, PATLDW is calculated for the weighted patent network.	Authors' elaboration on FP5 administrative database, DG RTD, Dir A and OECD REGPAT database
$FPGE_i$ / $PATGEW_i$	Global Embeddedness – the structure of the network behind a region's immediate neighborhood. FPGE is calculated for the binary FP network, PATGEW is calculated for the weighted patent network.	Authors' elaboration on FP5 administrative database, DG RTD, Dir A and OECD REGPAT database
$FPENQ_i$ / $FPENQW_i$ $PATENQ_i$ / $PATENQW_i$	Ego Network Quality – a comprehensive measure of network position. FPENQ is calculated for the binary FP network, FPENQW is calculated for the weighted FP network, PATENQ is calculated for the binary patent network and PATENQW is calculated for the weighted patent network.	Authors' elaboration on FP5 administrative database, DG RTD, Dir A and OECD REGPAT database
$FPDEG_i$ / $PATDEG_i$	The number of a region's direct partners in the network. FPDEG is calculated for the binary FP network and PATDEG is calculated for the binary patent network.	Authors' elaboration on FP5 administrative database, DG RTD, Dir A and OECD REGPAT database
$PATSTCKN_i$	National patent stock corresponding to the given region	Authors' elaboration on Eurostat NewCronos
$PUBSTCKN_i$	National publication stock corresponding to the given region	Authors' elaboration on Eurostat NewCronos
$AGGL_i$	Index of agglomeration. Size-adjusted location quotient of employment in technology- and knowledge-intensive sectors: high and medium high technology manufacturing, high technology services, knowledge intensive market services, financial services, amenity services – health, education, recreation. For more details see Varga, Pontikakis and Chorafakis (2012)	Authors elaboration of Eurostat NewCronos

Table 2. Variable descriptive statistics

	PUB	*PAT*	*GRD*	*PUBSTCKN*	*PATSTCKN*	
N	189	189	189	189	189	
Mean	2000.28	314.55	730.30	15 533.05	30 105.69	
Std.dev.	2576.39	519.94	1 212.11	14 929.01	36 317.06	
Min	3	0.01	1	25	11	
Max	21 050	3 282.27	11 314	41 111	98 481	
	FPDEG	*FPENQ*	*FPENQW*	*FPGE*	*FPLD*	*FPKP*
N	189	189	189	189	189	189
Mean	131.54	7 759 859	132 352	657 315	7 973	996
Std.dev.	42.58	1 599 143	18 046	934 644	36 154	252
Min	8	3 852 354	58 986	957	36	681
Max	186	9 300 660	156 383	4 378 189	12 388	2 105
	PATDEG	*PATENQ*	*PATENQW*	*PATGEW*	*PATLDW*	*PATKPW*
N	189	189	189	189	189	189
Mean	52.53	4 995 496	30 350	7 310	8.905	12.766
Std.dev.	36.16	1 544 242	11 459	6 422	7.125	20.355
Min	0	0	0	0	0.0002	0.008
Max	133	8 066 975	48 263	20 607	20.717	130.107

Following Romer (1990), the importance of knowledge stocks (or a 'standing on the shoulders of giants' effect) for knowledge production has been verified empirically (Furman, Porter and Stern, 2002; Zucker et al. 2007). In order to capture this effect we use proxies of national knowledge stocks (available for all regions in the given country) corresponding to the two different knowledge types ($PUBSTCKN_i$, $PATSTCKN_i$). For technological knowledge we use national patent stock and the national publication stock for scientific knowledge. Patent stock is calculated according to the perpetual inventory method for the 1992–1998 period (see the details in Varga, Pontikakis and Chorafakis 2012) while publication stock is a simple sum of the count of publications in the period of 2000-2002. Variable description is in Table 1, while descriptive statistics of the main variables are presented in Table 2.

5. Interregional knowledge network quality and research productivity of European regions: Empirical analysis

In this section we present estimation results for equation (10). Knowledge production in scientific research (resulting in publications) and technological inventions (measured by patents) will be investigated for 189 EU NUTS 2 regions. Cross-sectional econometric estimations for parameter β_1 will receive a particular attention as this parameter proxies interregional ego network effects on regional research productivity. Additional to core model variables we control for local knowledge flow impacts estimated by the parameter of the variable AGGL measuring agglomeration of knowledge intensive industries in the region.[10] We also test for spatial dependence in order to control for interregional knowledge flows communicated by channels different from FP collaborations or co-patenting. In this respect the significant parameter of the spatial lag variable is taken as a sign of the role of such interregional knowledge flows.[11] Due to the presence of the interaction term in equation (11) multicollinearity is a potential problem. We test for it by the Multicollinearity Condition Number (MCN).[12]

It has to be emphasized at this point that in our analysis the impacts of three types of knowledge flows are taken into account, one a-spatial and two spatial. A-spatial knowledge flows are communicated through (FP or co-patent) collaborations and their impacts are represented by the parameter of the ENQ index. These flows are called a-spatial since we assume that learning through these networks does not necessarily require the actors to live in spatial proximity to each other. Therefore spatial distance between the regions in these networks is not taken into account. The only issue we are interested in is the extent to which regional research productivity is influenced by the knowledge potentially accessed from collaboration partners located in different regions. Partners might be situated thousands of kilometers away or in a proximate region they are represented in the network as direct collaborators. They are thus neighbors in the network but not necessarily neighbors in the spatial sense.

However some of the knowledge flows might require spatial proximity of actors when the speed of information communication or tacit elements of knowledge require frequent face-to-face connections. In our regression context we also account for the impacts of these types of knowledge flows. The coefficient of the variable AGGL is assumed to proxy intraregional knowledge transfers whereas the coefficient of the spatially lagged dependent variable is expected to capture learning between actors located in proximate regions. While accounting for spatial knowledge flows in the models we assume (and test) whether spatial proximity is indeed important for certain

---

[10] Following Varga, Pontikakis and Chorafakis (2012) the index is a size-adjusted (in the spirit of the index developed by Elison and Glaeser 1997) variation of the popular location quotient (LQ) measure and is calculated as:

$AGGL_i = [(EMPKI_i / EMPKI_{EU}) / (EMP_i / EMP_{EU})] / [1 - \sum_j (EMPKI_{i,j} / EMPKI_{j,EU})]*[1 - (EMP_i / EMP_{EU})]$,

where $EMPKI_j$ and EMPKI are employment in knowledge intensive economic sector j and the total of knowledge intensive sectors, EMP is total employment and the subscripts *i* and *EU* stand for region and EU aggregate respectively. A significant and positive parameter of AGGL indicates a positive relation between knowledge output (publications or patents) and the agglomeration of knowledge intensive industries usually found instrumental in innovation such as high and medium technology manufacturing and business services. As common in KPF studies we interpret this result as a sign of influential knowledge flows from the local knowledge intensive industry to the production of new knowledge.

[11] The general expression for the spatial lag model is

$$y = \rho Wy + x\beta + \varepsilon,$$

where y is an N by 1 vector of dependent observations, Wy is an N by 1 vector of lagged dependent observations, ρ is a spatial autoregressive parameter, x is an N by K matrix of exogenous explanatory variables, β is a K by 1 vector of respective coefficients, and ε is an N by 1 vector of independent disturbance terms. Because the spatially lagged dependent term is correlated with the errors and as such endogenous, the OLS estimator is biased and inconsistent. Instead of OLS, other estimation methods such as Maximum Likelihood, Instrumental Variables or General Methods of Moments must be applied to the spatial lag model (Anselin 1988).

[12] The value of MCN exceeding 30 suggests a potential problem of specification (Belsley, Kuh, and Welsch 1980).

Table 3. Regression Results for Log (Patents) for 189 EU regions, 2002 (N=189)

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	ML-Spatial Lag (NEIGH)
Constant	-1.972***	-1.642***	-1.501***	0.101	0.195	-0.769**	-2.071***	-0.837	-0.791
	(0.335)	(0.353)	(0.333)	(0.433)	(0.500)	(0.375)	(0.463)	(0.593)	(0.566)
W_Log(PAT)									0.027***
									(0.008)
Log(GRD(-2))	1.123***	1.044***	0.876***	0.899***	0.890***	0.602***	0.638***	0.497***	0.526***
	(0.058)	(0.064)	(0.076)	(0.062)	(0.078)	(0.106)	(0.101)	(0.108)	(0.104)
Log(GRD(-2))* PATKPW		0.224***			0.081				
		(0.086)			(0.084)				
Log(GRD(-2))* PATLCW			1.096***		-0.092				
			(0.233)		(0.373)				
Log(GRD(-2))* PATGEW				-0.504***	-0.507***				
				(0.076)	(0.132)				
Log(GRD(-2))* PATENQW						0.456***	0.294***	0.257***	0.195**
						(0.080)	(0.085)	(0.084)	(0.082)
Log(PATSTCKN(-2))							0.192***	0.171***	0.114***
							(0.043)	(0.043)	(0.043)
Log(AGGL(-2))								1.195***	1.149***
								(0.372)	(0.355)
R^2-adj	0.67	0.68	0.70	0.73	0.73	0.72	0.74	0.76	0.77
LIK									-269
Multicollinearity Condition Number	7.4	8.8	11.0	11.5	14.1	15.7	19.5	22.2	22.2
LM-Err									
Neigh	30.46***	21.36***	14.14***	8.56***	7.84***	12.41***	9.03***	5.67**	
INV1	59.21***	34.30***	22.54***	9.13***	7.31***	12.19***	5.38**	2.33	
INV2	16.60***	9.96***	6.54***	2.75*	2.22	4.01**	1.69	0.98	
LM-Lag									
Neigh	42.13***	35.01***	28.90***	20.07***	3.19*	23.39***	12.65***	12.65***	
INV1	48.06***	39.82***	30.86***	18.51***	18.90***	21.25***	10.94***	7.09***	
INV2	17.04***	13.01***	9.89***	3.75**	17.24***	4.64***	3.08*	1.23***	
LR-Lag									12.349***
LM-Err									
Neigh									1.456
INV1									0.313
INV2									0364

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; W_Log(PAT) is the spatially lagged dependent variable where W stands for the weights matrix INV2. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

knowledge elements to be communicated. Note that while specific knowledge communication channels are not distinguished when we test for their spatially mediated impacts, in the a-spatial cases the communication channels are explicitly taken into account (FP collaboration or co-patenting).

Table 3 shows the results on regional patenting. The highly significant and positive parameter of the research variable and the good regression fit in the second column are usual findings in the KPF literature on patenting. Models 2 to 4 test the separate impacts of the three dimensions of ENQ on research productivity. Impacts with binary and weighted patent networks were both estimated but since estimations for the latter network resulted in somewhat better fits to the data, the weighted network results are presented here. Knowledge potential (PATKPW) and the intensity of collaborations (PATLCW) in the ego network are both positively related with research productivity as signaled by the highly significant coefficients.

The estimated parameter of global embeddedness (PATGEW) is significant but negative. It is the particular structure of the entire co-patenting network that lies behind this somewhat surprising finding of negative association between global knowledge access and regional R&D productivity. Data clearly tell us that the

European co-patenting network follows a strong core-periphery structure. High patenting regions in the core tend to collaborate with each other frequently while low patenting regions in the periphery have few ties to the core.[13] Since most of the knowledge agglomerates in the core, additional knowledge that core regions can learn from the periphery is only marginal. On the other hand, peripheral regions, due to their rare connections to the core, do not find considerable knowledge in their direct neighborhood but farther away in the network (which is captured by GE) there is important knowledge, namely that of the core regions.[14] We decided to leave this result as it is. Our data do not allow for a more detailed analysis at this point. Information on specific technologies over a longer time frame will certainly provide a deeper knowledge about the role of global embeddedness in regional R&D productivity. With such data at hand, experiments with different weighting methods in ENQ (equation 1) could also become possible.

Overall considered, it is clear from Table 3 that accounting for the impact of the three sub-indices (measuring the three dimensions of ego network quality) increases regression fit. Additionally, differences in the estimated values of β_0 and β_1 in Models 2 to 4 suggest that the three sub-indices indeed capture different dimensions of ego network quality. Furthermore, the results of Model 5 in Table 3 provide an additional, practical argument for the development of a comprehensive measure of network characteristics. When the three sub-indices are all included in the equation at the same time cross-correlations among them result in insignificant parameter estimates and strangely changed coefficient signs. However, it does not mean that there is no solution for testing the parallel impacts in a regression context. When the parallel impacts of the three sub-indices are accounted for by the comprehensive ENQ index no such statistical problems are expected.

ENQ enters the regression with a strongly significant and positive parameter in Model 5 resulting in an equation with R-squared 7.5 percent higher than without considering the effects of interregional networks on R&D productivity in Model 1. Repeating earlier results in the literature (e.g., Varga, Pontikakis an Chorafakis 2012) it is found that both national knowledge stock and agglomeration of knowledge intensive industries affect regional patenting positively. The positive and highly significant parameter of the spatially lagged dependent variable suggests that knowledge flows between spatially proximate regions are still important sources of invention even after

[13] Note the similarity between the structure of the empirical patent network and that of the theoretical scalefree network in Figure 2. Though to a somewhat lesser extent but a similar pattern exists in the FP5 network, which explains the comparable findings on GE impacts for publication research productivity (see Table 6 for further details).

[14] As a proximate measure of the relative size of knowledge accessed from outside the ego network we calculated the share of GE over ENQ for each region in the sample. Core regions yield extremely low values (e.g., for Ile de France it is below 1 percent) while on the perihpery the share of globally accessible knowledge above 90 percent is not an exception. This suggests that for several regions in the periphery globally available knowledge can be about nine times higher than the knowledge accessible from their individual networks. We also experimented with different methods to separate core and peripheral regions empirically. Estimation results (not reported here) suggest that the core and the perihpery indeed follow different patterns in utilizing global knowledge in generating new technologies.

Table 4. Regression Results for Log (Publications) for 189 EU regions, 2002 (N=189)

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	ML-Spatial Lag (INV1)
Constant	1.402*** (0.229)	2.419*** (0.262)	2.658*** (0.285)	2.554*** (0.280)	2.330*** (0.311)	2.584*** (0.288)	2.039*** (0.517)	1.980*** (0.402)	1.859*** (0.389)
W_Log(PUB)									-0.004*** (0.001)
Log(GRD(-2))	0.941*** (0.039)	0.206* (0.120)	0.301*** (0105)	0.794*** (0.043)	-1.641** (0.819)	0.462*** (0.088)	0.366*** (0.114)	0.377*** (0.095)	0.459*** (0.095)
Log(GRD(-2))* FPKP		0.612*** (0.096)			1.614** (0.643)				
Log(GRD(-2))* FPLC			0.489*** (0.076)		0.884*** (0.339)				
Log(GRD(-2))* FPGE				-0.536*** (0.087)	1.873** (0.806)				
Log(GRD(-2))* FPENQ						0.346*** (0.058)	0.380*** (0.060)	0.377*** (0.059)	0.334*** (0.058)
Log(PUBSTCKN(-2))							0.096** (0.045)	0.096** (0.045)	0.140*** (0.045)
Log(AGGL(-2))							0.049 (0.267		
R^2-adj	0.75	0.80	0.80	0.79	0.80	0.79	0.79	0.80	0.81
LIK									-209
Multicollinearity Condition Number	7.4	27.5	23.6	10.5	223.3	19.1	27.9	23.3	23.3
LM-Err									
Neigh	0.123	0.121	0.068	0.128	0.029	0.117	0.277	0.261	
INV1	1.213	1.081	0.399	0.206	0.165	0.359	1.144	1.070	
INV2	0.172	0.252	0.086	1.044	0.095	0.031	0.210	0.201	
LM-Lag									
Neigh	11.008***	7.041***	4.941**	6.330**	5.963**	4.461**	8.139***	8.164***	
INV1	12.988***	6.348**	4.901**	1.528	4.220**	5.506**	10.833***	9.704***	
INV2	5.106**	1.409	1.289	6.321**	0.937	1.977	2.820*	2..523	
LR-Lag									10.50***
LM-Err									
Neigh									0.001
INV1									0.002
INV2									0.267

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; W_Log(PAT) is the spatially lagged dependent variable where W stands for the weights matrix INV2. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

controlling for the impact of knowledge communicated through interregional (a-spatial) co-patenting networks.

Table 4 exhibits empirical results for regional R&D productivity in scientific publications. Models 1 to 6 in Table 4 follow the patterns already discussed for patent production: components of ENQ (this time with binary network data) when included separately in the equations enter the models with significant and varying parameters, and the parameter of ENQ is itself positive and significant while regression fit increases. National level scientific knowledge relates positively to publication output, but regional agglomeration of technology-intensive industries has no significant effect on the production of new scientific knowledge (Models 7 and 8). This finding, together with the significant network quality effect, suggests that success in scientific research is more related to international and national embeddedness than to industrial agglomeration. This result echoes what is a frequent observation: high quality research universities are not necessarily located in large cities. They can be equally successful in scientific knowledge production in smaller but well-connected geographical areas (e.g., Varga 2000). The significant but negative coefficient of the spatially lagged dependent variable in Model 9 underlines what is just observed about the role of agglomeration. Regions with high levels of scientific publication do not necessarily locate in spatial proximity, but they tend to scatter geographically. This latter finding underlines what is suggested by Models 7 and 8 in Table 4. Knowledge inputs to scientific research can successfully be transported over large distances due perhaps to the presence of institutional and cognitive proximities (Boschma 2005).

5. Spatial regimes in research performance: Do Objective 1 and EU 12 border regions follow different paths?

Though it is not possible to directly test the effects of co-patenting and FP network quality on EU neighboring countries we were able to get some estimate by testing the impacts of co-patenting and FP network quality on research productivity of NUTS 2 regions which possess the characteristics of EU neighboring countries in two respects: first, their GDP per capita is below the 75 % or EU average (Objective 1 regions) and second, they are located at the border of the old EU 12 territory (border regions).

Our empirical model is modified by replacing equations (8) and (9) by equations (11) and (12).

$$a_{1,i} = \beta_0 Log(RD_{i,\ t\text{-}k}) + \beta_1 Log(ENQ_i) + \beta_2 REGIMEDUMMY_i \quad (11)$$

$$\begin{aligned} Log(K_{i,\ t}) = \alpha_0 &+ \beta_0 Log(RD_{i,\ t\text{-}k}) + \beta_1 Log(ENQ_i)*Log(RD_{i,\ t\text{-}k}) + \\ &+ \beta_2\ REGIMEDUMMY_i*Log(RD_{i,\ t\text{-}k}) + \\ &+ \alpha_2 Log(KSTCK_{N,\ t\text{-}k}) + Z_i + \varepsilon_i. \end{aligned} \quad (12)$$

The change in equations (11) and (12) as compared to equations (9) and (10) is that now a dummy variable stands for a presumed spatial regime effect (i.e., Objective 1 or border region effects). REGIMEDUMMY may stand for a dummy for Objective 1 regions (OBJ1), a dummy for regions at the boundary of the EU12 territory from the EU-side (WEST) or for a dummy for border regions at the non-EU side (EAST).

Table 5 shows spatial lag estimation results for patenting. For comparison, outcomes of a previously reported final estimation are depicted in the 2nd column (we call this model here as the General model). Inclusion of the WEST and EAST border dummies do not change estimation results meaningfully while the parameter of WEST is insignificant and that of the EAST variable is only marginally significant (4th and 5th columns). However, from the second model it is clear that Objective 1 regions behave differently. The interaction variable with the Obj1 dummy is highly significant and at the same time regression fit increases in comparison to the General model. The significant negative parameter suggests that even with a same ENQ an Objective 1 region reaches a lower productivity than a non-Objective 1 region. This difference might be due to lack of absorptive capacities on the side of firms or a less developed regional innovation infrastructure.

Results for publication in Table 6 are markedly different from what is found for patenting. One of the differences is that the EAST and WEST dummies enter the model with insignificant parameters. Additionally after comparing the 2nd and the 3rd columns in Table 6 it becomes clear that parameter estimates do not change meaningfully and the interaction term with the Obj1 dummy is highly significant and

Table 5. Are Objective 1 and Border Regions Different in Patenting? Regression Results for Log (Patents) for 189 EU regions, 2002 (N=189)

Model	General	Objective 1	West	East
Estimation	ML- Spatial Lag (NEIGH)	ML- Spatial Lag (NEIGH)	ML- Spatial Lag (NEIGH)	ML- Spatial Lag (NEIGH)
Constant	-0.791	-0.657	-0.778	-0.458
	(0.566)	(0.540)	(0.566)	(0.596)
W_Log(PAT)	0.027***	0.022***	0.027***	0.030***
	(0.008)	(0.007)	(0.007)	(0.008)
Log(GRD(-2))	0.526***	0.616***	0.535***	0.506***
	(0.104)	(0.101)	(0.104)	(0.103)
Log(GRD(-2))* PATENQW	0.195**	0.168**	0.184**	0.190**
	(0.082)	(0.078)	(0.083)	(0.081)
Log(PATSTCKN(-2))	0.114***	0.086**	0.110**	0.088*
	(0.043)	(0.042)	(0.044)	(0.046)
Log(AGGL(-2))	1.149***	0.726**	1.173***	1.277***
	(0.355)	(0.351)	(0.357)	(0.361)
Log(GRD(-2))* OBJ1		-0.147***		
		(0.033)		
Log(GRD(-2))* WEST			0.038	
			(0.058)	
Log(GRD(-2))* EAST				-0.207*
				(0.122)
R^2-adj	0.77	0.80	0.77	0.78
LIK	-269	-260	-269	-268
Multicollinearity Condition Number	22.2	23.1	18.0	17.7
LR-Lag	12.349***	9.443***	12.544***	14.637***
LM-Err				
Neigh	1.456	0.177	1.721	1.293
INV1	0.313	0.048	0.405	0.261
INV2	0364	0.020	0.470	0.346

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; W_Log(PAT) is the spatially lagged dependent variable where W stands for the weights matrix INV2. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

Table 6. Are Objective 1 and Border Regions Different in Publication? Regression Results for Log (Publications) for 189 EU regions, 2002 (N=189)

Model	General	Objective 1	West	East
Estimation	ML-Spatial Lag (INV1)	ML-Spatial Lag (NEIGH)	ML-Spatial Lag (INV1)	ML-Spatial Lag (INV1)
Constant	1.859***	1.230***	1.816***	1.767***
	(0.389)	(0.414)	(0.393)	(0.399)
W_Log(PUB)	-0.004***	-0.012***	-0.003***	-0.004***
	(0.001)	(0.004)	(0.001)	(0.001)
Log(GRD(-2))	0.459***	0.468***	0.464***	0.461***
	(0.095)	(0.093)	(0.095)	(0.095)
Log(GRD(-2))* FPENQ	0.334***	0.338***	0.329***	0.335***
	(0.058)	(0.056)	(0.058)	(0.058)
Log(PUBSTCKN(-2))	0.140***	0.160***	0.143***	0.150***
	(0.045)	(0.045)	(0.046)	(0.047)
Log(GRD(-2))* OBJ1		0.075***		
		(0.023)		
Log(GRD(-2))* WEST			-0.03	
			(0.041)	
Log(GRD(-2))* EAST				0.082
				(0.085)
R^2-adj	0.81	0.82	0.81	0.81
LIK	-209	-205	-209	-209
Multicollinearity Condition Number	23.3	24.3	23.6	23.4
LR-Lag	10.50***	8.41***	10.20***	11.28***
LM-Err				
Neigh	0.001	0.004	0.010	0.000
INV1	0.002	0.047	0.222	0.392
INV2	0.267	0.011	0.003	0.003

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; W_Log(PAT) is the spatially lagged dependent variable where W stands for the weights matrix INV2. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

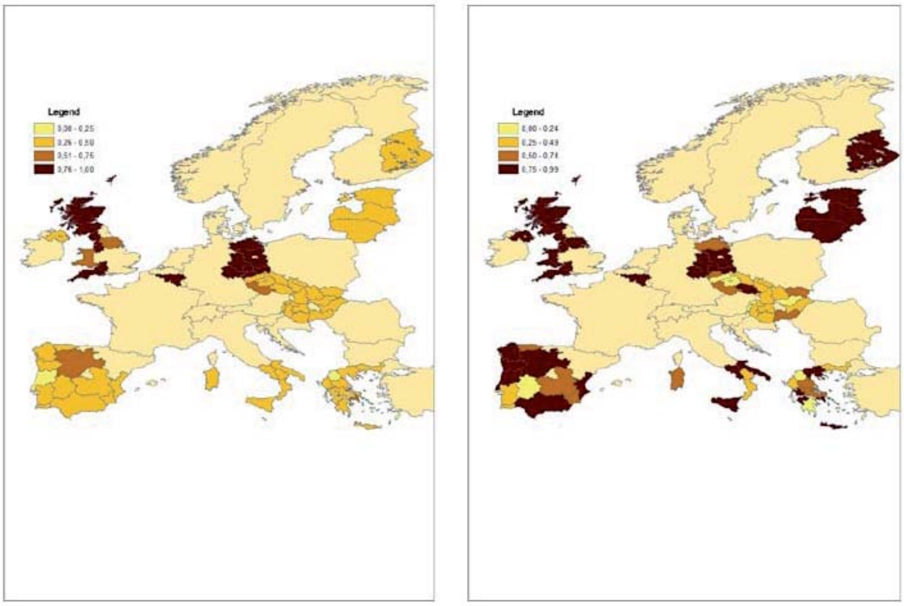

Figure 1. Objective 1 regions in the sample: ENQ in patenting (left panel) and publications (right panel)

positive. The coefficient of the variable is small suggesting that research productivity in publication is slightly higher for Objective 1 regions in comparison to the European average.

Figure 1 depicts ENQ values in patenting and publication for Objective 1 regions included in our sample. Considering the focus of the SEARCH project a particularly interesting set of regions are situated in Central and Eastern European countries that joined the EU in the last waves of accession. We assume that in many respect these regions share characteristics with several of the regions in EU neighboring countries. Our data set include East Germany, Czech and Slovak Republic and Hungary. With the exception of the capital regions the rest of the territories of these countries belong to the Objective 1 category. The left panel in Figure 1 clearly indicates that co-patenting network values measured by ENQ in most of the CEE regions in the sample are below the median (0.50). For publications (right panel of Figure 1) the situation is somewhat more advantageous for CEE countries as some of their regions with high quality universities succeed to build decent FP networks resulting in higher-than-median ENQ values.

Table 8 reports average ENQ and research productivity values for Objective 1 and rest of EU regions together with some input and output variables in knowledge production. R&D expenditures of Objective 1 regions in our sample are about one-quarter of that of the non-Objective 1 EU regions on average. Utilization of this markedly lower research input in Objective 1 regions is less successful in patenting than in scientific publications. While the average number of patents originating from Objective 1 regions is about 15 % of the corresponding value observed in rest of EU regions, the average number of publications is half of what is produced in non

Objective 1 regions. The figures in Table 8 draw attention to the role of extra regional knowledge network qualities in these differences. While the quality of co-patenting

Table 8. Objective 1 regions in comparison with Rest of EU: Average values of outputs and R&D input of knowledge production, network quality and research productivity

	PUB	PAT	GRD (M EUR)	ENQ_FP	ENQ_PAT	RD productivity PUB	RD productivity PAT
Obj 1 regions	1 197	64	270	0.63	0.47	0.76	0.55
Rest of EU	2 431	449	977	0.76	0.71	0.73	0.74
Whole sample	2 000	315	730	0.72	0.63	0.70	0.65

networks in Objective 1 regions is about two-third of what is observed in rest of EU regions, for publication the lag behind more advanced EU regions is less pronounced. An average Objective 1 region's ENQ in the FP network is about 80 percent of what is calculated for rest of EU regions. These differences are then reflected in differences in the corresponding research productivity values. While a one percentage increase in R&D expenditures in Objective 1 regions results in an increase of patenting which is about three-fourth of what is observed in rest of EU regions research productivities in publications in the two types of regions are about the same – even a bit higher for Objective 1 regions (0.76 for Objective 1 regions and 0.73 for rest of Europe). Since Objective 1 regions in CEE countries exhibit lower ENQ values in both patenting and publication than the rest of Objective 1 regions (as depicted in Figure 1) corresponding R&D productivities in CEE regions are also somewhat weaker suggesting even lower values for EU neighboring country regions.

6. Summary and conclusions

Our paper contributes to the emerging literature on the role of interregional knowledge flows in the regional production of new (scientific and technological) knowledge. An especially promising tool of research in this area is network analysis, which is applied in our study as well. Network analytic tools have been increasingly employed in studying the flows of knowledge in two, more or less separately developed scientific literatures. "A-spatial" approaches mostly appearing in the science and technology literature study the impacts of different characteristics such as size, centrality, density, tie strength or knowledge diversity of collaboration networks among firms, research institutions and alike. The influence of different network characteristics are examined individually and the selection of particular network features studied usually related to actual research questions and do not seem to follow an underlying theoretical agreement in the field.

Several of the "spatial" studies appearing in the regional economics and economic geography literature show a different origin and a somewhat different interest. Their empirical approach is significantly influenced by some spatial econometric techniques introduced in early studies searching for the spatial extent of knowledge spillovers (Anselin, Varga and Acs 1997). In these studies the impact of R&D spillovers originated from different geographical locations are modeled via spatially lagged

R&D variables. Significant parameters of these variables are understood as indicators of the impact of R&D carried out in a distant spatial unit on knowledge produced in the region.

It seems that a solution first applied in the early studies to estimate the spatial boundaries of R&D spillovers became the source of inspiration for the way interregional knowledge networks are modeled. Many of the spatial network studies apply this intuition by replacing spatial weights matrices with matrices representing interregional collaborations. With this technique it became possible to study the impact of R&D carried out by partners in a different spatial unit on knowledge produced in the region. Different weighting schemes are applied in this literature. When no weights are implemented then regional knowledge production is related to the total of R&D in partner regions (Maggioni and Uberti 2011, Varga, Pontikakis and Chorafakis 2012). Alternatively, the impacts of weighted averages (with number of collaborative projects as weights such as in Ponds, van Oort and Frenken 2012) of R&D in partner regions can be studied.

The present paper is an attempt to integrate techniques mainly applied in a-spatial studies with solutions implemented in spatial analyses. Following the theory of innovation we developed a systematic scheme for weighting R&D in partner regions with network features frequently appearing in several (mostly non-spatial) studies (tie strength, number of edges, density of interactions, network distance, knowledge diversity). The resulting ENQ index is taken as a measure of the level of knowledge accessible from an interregional network. Thus the interest behind ENQ is the same as in the spatial studies (i.e., the impact of R&D in partner regions). The difference is in the broader set of network features that we take into account in the analysis.

To structure the problem of interregional knowledge network effects on research productivity at the regional level we directed attention to the quality of knowledge that can be accessed from a particular network. To measure such impact we introduced the index of ego network quality (ENQ). ENQ summarizes three features of networks: the knowledge already accumulated by immediate network partners (Knowledge Potential – KP), the frequency of collaborations among immediate network partners (Local Connectivity – LC) and the region's embeddeddness in the entire knowledge network through the connections of their direct and indirect network partners (Global Embeddeddness – GE).

To test the validity of the new network measure, a systematic spatial econometric analysis was carried out with European regional data. Specifically, the role of the quality of collaboration networks in research addressing the development of technological inventions (measured by patents) on the one hand and scientific publications on the other were analyzed. We found that the quality of interregional networks in both areas of knowledge production is indeed a significant contributor to R&D productivity. We also found that the pure number of collaborations, which is the most frequently used variable in spatial network studies is not a suitable proxy of interregional network effects in R&D productivity contrary to ENQ. Our results show that a more comprehensive approach taking into account several local and global features of the network surrounding the given region provides better insights into the network effects in regional knowledge production.

Empirical analyses in this paper resulted in several additional interesting observations on the role of space in different types of knowledge generations. The finding, that non-spatially mediated learning in (FP or patent) collaboration networks significantly enhances research productivity, comes up strongly from the regression results for both types of knowledge creation. However, spatial proximity plays different roles in patenting and scientific publication. While local agglomeration of knowledge industries affect patenting positively together with patenting carried out in proximate regions, spatially mediated knowledge flows do not seem to be significant sources of science-oriented research resulting in publications. On the contrary, the significantly negative parameter of the spatially lagged publication variable suggests a "chessboard-like" spatial arrangement of scientific research institutions. They do not tend to spatially agglomerate but it also seems that their success in science do not necessarily need that agglomeration.

We were also able to get some estimates as to the likely situation in EU neighboring country regions by testing the impacts of co-patenting and FP network quality on research productivity in Objective 1 NUTS 2 regions. Our analysis yielded interesting findings. While the quality of co-patenting networks in Objective 1 regions is about two-third of that of in rest of EU regions, for publication the lag behind more advanced EU regions is less pronounced. Objective 1 region's ENQ is about 80 percent of what is calculated for rest of EU regions. These differences are then reflected in differences in the corresponding research productivity values.

These latter observations suggest that those EU neighboring country regions where good quality universities and public research institutions are located could potentially build research collaboration networks competitive with networks maintained by many European regions. Thus intensifying the participation of neighboring countries in EU Framework Program funded research projects could result in an increase in research productivity of those neighboring country regions where already substantial research capacities are built at local universities or public research facilities. This increased research productivity might later form the basis of regional economic development policies taking advantage of potentials accumulated at their higher education institutions or public research institutes. Policies aiming at attracting private research labs of industries closely related to the region's research specialization paired with suitable complementary interventions (such as building up human capital assets in the region or physical infrastructure development) could potentially initiate a longer run cumulative process that possibly ends up in a substantial regional industry concentration (Varga, Pontikakis and Chorafakis 2013).

References

Acs, Z., Anselin, L., and Varga, A. (2002) Patents and innovation counts as measures of regional production of new knowledge. *Research Policy* 31, 1069-1085.

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

Barabási, A., and Albert, R. (1999) Emergence of scaling in random networks. *Science 286* (5439), 509-512.

Belsley, D., Kuh, E., and Welsch, R. (1980): *Regression Diagnostics, Identifying Influential Data and Sources of Collinearity*. New York: Willey.

Bonacich, P. (1972) Factoring and weighting approaches to clique identification. *Journal of Mathematical Sociology* 2, 113-120.

Bonacich, P. (2007) Some unique properties of eigenvector centrality. *Social Networks* 29, 555-564.

Boschma, R. (2005) Proximity and innovation: A critical assessment. *Regional Studies* 39, 61-74.

Breschi, S. and Lenzi, C. (2011) *Net and the city. Co-invention networks and the inventive productivity of US cities*. mimeo.

Breschi, S. and Lissoni, F. (2009). Mobility of inventors and networks of collaboration: An anatomy of localised knowledge flows. *Journal of Economic Geography* 9, 439-468.

Broekel, T., Buerger, M., Brenner, M. (2010) An investigation of the relation between cooperation and the innovative success of German regions. *Papers in Evolutionary Economic Geography (PEEG)* 1011, Utrecht University, Section of Economic Geography, revised Oct 2010.

Burt, R.S. (1992): *Structural Holes*. Harvard University Press. Cambridge, MA.

Burt, R.S., Hogarth, R.M., Michaud C. (2000): The Social Capital of French and American Managers. *Organization Science* 11:123-147.

Burton, P., Yu W., Prybutok, V. (2010): Social network position and its relationship to performance of IT professionals.(Report). Informing Science: the International Journal of an Emerging Transdiscipline. Informing Science Institute. *HighBeam Research*

Cainelli, G., Maggioni, M., Uberti, E and De Felice, A. (2010) *The strength of strong ties: co-authorship and productivity among Italian economists*. 'Marco Fanno' working papers 125, Department of Economics, University of Padova.

Cantwell, J. and Iammarino, S. (2003) *Multinational corporations and European regional systems of innovation*. Routledge.

Coleman, J.S. (1986): Social Theory, Social Research, and a Theory of Action. *American Journal of Sociology*, 91: 1309-1335.

Cross, R., Cummings, J.N. (2004) Tie and Network Correlates of Individual Performance in Knowledge-Intensive Work. *The Academy of Management Journal*. 47 928-937.

Donckels, R. Lambrecht, J. (1997): The Network Position of Small Businesses: An Explanatory Model. *Journal of Small Business Management*, 35(2): 13-26.

Feldman, M. and Audretsch, D. (1999) Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review* 43, 409-429.

Fischer, M. and Varga, A. (2002) Technological innovation and interfirm cooperation. An exploratory analysis using survey data from manufacturing firms in the metropolitan region of Vienna. *International Journal of Technology Management* 24, 724-742.

Foray, D., David, P. and Hall, B. (2009) *Smart Specialisation – The Concept*. Knowledge Economists Policy Brief No 9, June 2009.

Freeman, L.C. (1979) Centrality in social networks: I. Conceptual clarification. S*ocial Networks*. 1, 215-239.

Frenken, K., van Oort F. and Verburg, T. (2007) Related variety, unrelated variety and regional economic growth. *Regional Studies* 41(5): 685-697.

Fujita M and Thisse J (2002) *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth.* Cambridge University Press Cambridge, MA, London, England.

Furman, J. L., Porter, M. E. and Stern, S. (2002) The determinants of national innovative capacity. *Research Policy* 31, 899-933

Ghinamo M. (2012) Explaining the variation in the empirical estimates of academic knowledge spillovers. *Journal of Regional Science* 52, 606–634.

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

Griliches Z. (1990) Patent statistics as economic indicators: a survey. *Journal of Economic Literature* 20, 1661-1707.

Hoekman, J., Frenken, K. and van Oort, F. (2009) Collaboration networks as carriers of knowledge spillovers: Evidence from EU27 regions. *Annals of Regional Science* 43, 721-738.

Hopp, W.J., Iravani, S., Liu, F., Stringer, M.J. (2010) *The Impact of Discussion, Awareness, and Collaboration Network Position on Research Performance of Engineering School Faculty*. Ross School of Business Paper No. 1164.

Jaffe, A. (1989) Real effects of academic research, *American Economic Review 79*: 957-970.

Jaffe, A., Trajtenberg, M., and Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics 108*: 577–598.

Jones, C. (1995) R&D-based models of economic growth, *Journal of Political Economy 103*(4): 759-784.

Koschatzky, K. (2000) *The regionalization of innovation policy in Germany – Theoretical foundations and recent evidence*, Working Papers Firms and Regions No. R1/2000, Fraunhofer Institute for Systems and Innovation Research (ISI), Department "Innovation Services and Regional Development".

Kretschmer, H. (2004). Author productivity and geodesic distance in bibliographic co-authorship networks , and visibility on the Web. *Scientometrics*, 60(3): 409-420.

Lundvall, B.A. (1992) *National Systems of Innovation*. Pinter Publishers, London.

Maggioni, M. and Uberti, T. (2011) Networks and geography in the economics of knowledge flows. *Quality and Quantity*, 1031-1051.

Maggioni, M., Nosvelli, M. and Uberti, E. (2006) Space Vs. Networks in the Geography of Innovation: A European Analysis. *Papers in Regional Science* 86, 471-493.

Maier, G., Kurka, B. and Trippl, M. (2007) *Knowledge Spillover Agents and Regional Development: Spatial Distribution and Mobility of Star Scientists*. WP 17/2007, Wi t Wien.

, E. and Moreno, R. (2012) *Skilled Labour Mobility, Networks and Knowledge Creation in Regions: A Panel Data Approach*. Unpublished manuscript.

, J. (2009) *Inventors on the move: tracing invertors' mobility and its spatial distribution*. Research Institute of Applied Economics Working Papers 2009/16, University of Barcelona.

Nelson, R.R. (ed.) (1993) *National Innovation Systems: A Comparative Analysis*. Oxford University Press, Oxford.

OECD (2009) REGPAT Database, October 2009.

Ponds, R. van Oort, F. and Frenken, K. (2009) Internationalization and regional embedding of scientific research in the Netherlands. In Varga, A. (Ed.) *Universities, Knowledge Transfer and Regional Development: Geography, Entrepreneurship and Policy*. Edward Elgar Publishers, 109–137.

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, W.W., Koput, K.W., Smith-Doerr, L., Owen-Smith, J. (1999) Network Position and Firm Performance: Organizational Returns to Collaboration in the Biotechnology Industry. In Andrews, S. B., Knoke, D. (eds.): *Networks In and Around Organizations*. JAI Press, Greenwich, CT.

Diez, R. (2002) Metropolitan Innovation Systems - A comparison between Barcelona, Stockholm, and Vienna. *International Regional Science Review* 25, 63-85.

Romer, P. M. (1990) Endogenous technological change, *Journal of Political Economy 5*(98): S71-S102.

Rumsey-Wairepo, A. (2006) *The association between co-authorship network structures and successful academic publishing among higher education scholars*. Brigham Young University.

Salmenkaita, J. P. (2004) Intangible capital in industrial research: Effects of network position on individual inventive productivity. In Bettis, R. (Ed.) *Strategy in transition*. Blackwell Publishing, 220-248, Malden, MA.

Schiller, D. and Diez, J. (2008) *Mobile star scientists as regional knowledge spillover agents*. IAREG Working Paper WP2/7.

Sebestyén, T. and Varga, A. (2013) Research productivity and the quality of interregional knowledge networks. *Annals of Regional Science* (forthcoming).

Tsai, W. (2001): Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance. *The Academy of Management Journal*, 44(5): 996-1004.

Van Der Deijl, H., Kelchtermans, S., Veugelers, R. (2011) *Researcher networks and productivity*. Paper presented at the DIME-DRUID ACADEMY Winter Conference 2011.

Van Raan, A. F. J. (2004) Measuring science. In Moed, H. F., Glänzel, W. and Schmoch, U. (Eds) *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*. Kluwer Academic Publishers, 19-50, Dordrecht.

Varga, A. (2000) Local academic knowledge transfers and the concentration of economic activity, *Journal of Regional Science 40*(2): 289-309.

Varga, A. (2006) The spatial dimension of innovation and growth: empirical research methodology and policy analysis, *European Planning Studies 9*: 1171-1186.

Varga, A. and Parag, A. (2009) Academic knowledge transfers and the structure of international research networks. In Varga, A. (Ed.) *Universities, Knowledge Transfer and Regional Development: Geography, Entrepreneurship and Policy*. Edward Elgar Publishers, 138-159.

Varga, A., Pontikakis, D. and Chorafakis, G. (2013) Metropolitan Edison and cosmopolitan Pasteur? Agglomeration and interregional research network effects on European R&D productivity. *Journal of Economic Geography* (forthcoming).

Watts, D.J., Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *Nature*, 393(6684): 409–410.

Zaheer, A., Bell, G.G. (2005) Benefiting from network position: firm capabilities, structural holes and performance. *Strategic Management Journal*, 26: 809-825.

Zucker L.G. Darby, M.R.,Furner, J., Liu R.C. and Ma, H. (2007) Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production, *Research Policy*, Vol. 36, No. 6, pp. 850-863.

AQR UB IREA

Utrecht University

SEED CENTER
SOUTH & EAST EUROPEAN DEVELOPMENT CENTER

CRENoS

LSE

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Brunel UNIVERSITY LONDON

Université Jean Monnet SAINT-ETIENNE

UNIVERSITAS QUINQUEECCLESIENSIS

Leibniz Universität Hannover

UNIVERSITY of TARTU

STATE UNIVERSITY HIGHER SCHOOL OF ECONOMICS

Cadi Ayyad University

ICBSS INTERNATIONAL CENTRE FOR BLACK SEA STUDIES

IEMed. Institut Europeu de la Mediterrània

האוניברסיטה העברית בירושלים
The Hebrew University of Jerusalem

TÜBİTAK