Are knowledge flows all alike?
Evidence from European regions

Francesco Quatraro*[a,c] and Stefano Usai[b]

a) University of Torino, Department of Economics and Statistics Cognetti de Martiis, Lungo Dora Siena 100a, 10128 Torino (Italy)
b) University of Cagliari, CRENoS, Via S. Ignazio, 78, Cagliari (Italy)
c) BRICK, Collegio Carlo Alberto, Moncalieri (Torino, Italy)

This is a pre-copyedited, author-produced version of an article accepted for publication in REGIONAL STUDIES, following peer review.
The Version of Record: Quatraro F, USAI, STEFANO (2017). Are knowledge flows all alike? Evidence from European regions. REGIONAL STUDIES, vol. 51, p. 1246-1258, ISSN: 0034-3404, is available online at https://doi.org/10.1080/00343404.2016.1240867

ABSTRACT. The paper investigates the impact of distance, contiguity and technological proximity on cross-regional knowledge flows, by comparing the evidence concerning co-inventorship, applicant-inventor relationships and citation flows. We find evidence of significant differences across these diverse kinds of knowledge flows for what concerns the role of distance, and the moderating role of contiguity and technological proximity. Moreover, we show that border effects may prove crucial in a twofold sense. On the one hand we show that contiguity between regions belonging to two different countries still plays a moderating role, although weaker as compared to that of within-country contiguity. On the other hand, regions sharing a frontier with a foreign country are more likely to exchange knowledge with this foreign country than other regions which are far away from the border.

Keywords: Knowledge Flows, Border regions, Patents, regional competitiveness, Europe, Gravity

JEL Classification Codes: R11, O33

Acknowledgments. Preliminary versions of this paper have been presented at the conference of the Italian branch of the Regional Science Association (AISRE) held in Palermo (September 2013); the Geography of Innovation conference held in Utrecht (January 2014); the RSA conference held in London (November 2013), at a seminar held at the University of Padua (November 2013) and at the annual meeting of the Italian Economists Society, held in Trento in October 2014. We thank the participants for comments and remarks. Francesco Quatraro acknowledges the funding of European Commission, within the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 266959 (PICK-ME).

1 Introduction

Since the seminal work by Jaffe et al. (1993), an increasing body of empirical literature has focused on the analysis of knowledge flows and spillovers, mainly drawing upon data about patent citations. Despite the wide range of empirical works, the debate about the localization of knowledge spillovers is still far from finding an exhaustive conclusion. A common criticism is that citations may prove to be a 'noisy' indicator of knowledge spillovers (Jaffe et al., 1998), since they do not always imply an actual flow of knowledge from cited to citing inventor. Indeed, Thompson and Fox Kean (2005) show that the results obtained by Jaffe et al. (1993) are due to an imperfect matching of patent data, which is likely to produce a biased evidence concerning the geographical clustering of citations. Following this result, Thompson (2006) proposes an alternative citing-cited patent matching scheme, showing that citations still appear to be localized both within and across international borders. More recently, Belenzon and Schankerman (2013) study the geography of university knowledge spillovers, confirming that citations to patents are localized and sensitive to border effects whilst citations to publications are not. Criscuolo and Verspagen (2008) extend the debate and the analysis to the European case and find that geographical distance is a factor that strongly diminishes the probability of knowledge flows. This probability is found to be influenced also by cognitive distance and time.

However, Jaffe et al. (1998), in light of a quantitative and qualitative analysis, conclude that geographic spillovers are underestimated by patent citations and point to the necessity to go beyond this indicator. This research avenue has been recently explored in some contributions (Picci, 2010, Maggioni et al. 2011, Capelli and Montobbio, 2013) which investigate other patent related indicators, such as collaborations among inventors and relationships among patent inventors and applicants.

Our paper intends to contribute exactly on these final suggestions by investigating the differential effects of proximity across different types of knowledge flows. We, therefore, extend the analysis of knowledge spillovers so as to consider cooperative relationships among inventors and their relationship with formal patent applicants (most often firms), besides citations as proxies of cross-regional knowledge flows. The paper's contribution to the field is three-fold. First, we compare three indicators of knowledge flows across regions in Europe in the last decade, *i*.e. citations, applicant-inventor links and co-inventorships, in order to ascertain if knowledge flows are all alike in terms of their dependence on geographical distance and contiguity. Secondly, we provide evidence of the moderating role of technological proximity on the effect of physical distance. Thirdly, we investigate the differential patterns of inter-national vs intra-national flows and knowledge exchanges among core and peripheral regions. Finally, we test the robustness of our results with respect to the role of proximities both along time and across sectors.

Our results indicate that these indicators show different responses to proximity, citations being less dependent on physical contiguity than co-inventorships. On the contrary, when one considers the role of technological proximity, citations appear to be more sensitive than co-inventorship. The applicant-inventor relationship always appears as an intermediate phenomenon. These different patterns can be explained by noticing that co-inventorships concern mainly the exchange of tacit knowledge, while citations are more likely to involve the flow of codified knowledge.

The paper is structured as follows. In section 2 we present and discuss our theoretical and empirical background. Section 3 describes the dataset, the variables and the methodology. In section 4 we present the results of econometric estimations, while in the final section we conclude with some policy implications.

2 Knowledge Flows, Proximity and Border Effects

According to the conventional Marshallian tradition (Meade, 1952; Viner, 1932), knowledge spillovers are qualified as ‘untraded’ interdependencies among firms. Knowledge spills over and engenders positive externalities essentially due to its non-exclusive and non-rival use (Arrow, 1962).

Systemic approaches to innovation activities depict the generation of technological knowledge as an outcome of a collective undertaking strongly influenced by the availability of external sources of knowledge and by the way in which interactions are organized and carried out (Allen, 1983; von Hippel, 1988; Lundvall, 1992; Nelson, 1993). Internal and external knowledge inputs are so complementary that too low levels of each of them can hinder the entire knowledge production process (Antonelli, 1999).

The collective and interactive dimension of technological knowledge raises the issue of proximity of innovating agents (Foray, 2004). A wide body of literature has shown that knowledge spillovers tend to be geographically clustered, and firms are likely to base their location choices on the opportunities of taking advantages of the positive feedbacks associated to co-location with other innovative actors (Audretsch and Feldman, 1996; Baptista and Swann, 1998). Feldman (1994a and 1994b) argues that co-location mitigates the inherent uncertainty of innovative activity: proximity enhances the ability of firms to exchange ideas and be aware of important incipient knowledge.

In this context, the distinction between tacit and codified knowledge is especially relevant. Definitions of tacit knowledge often recall the well-known Polanyi’s quotation according to which people know more than they can tell. In this sense, tacit knowledge is highly idiosyncratic and difficult to communicate. On the contrary, codified knowledge, thanks to a shared guidebook that allows for coding and decoding, is better transmittable and understandable by people knowing the codebook. However, knowledge is not created codified. Codification is indeed the outcome of a process triggered by intentional efforts of innovating agents. In this perspective codified and tacit knowledge are not to be considered as discrete states, but rather as two extreme poles of a continuum (Saviotti, 1998; Cowan, David and Foray, 2000). Von Hipple (1994) explains that most of economic agents' tacit knowledge is 'sticky'; *i*.e. highly contextual and uncertain and concludes that it is best transmitted via face-to-face interaction and through frequent and repeated contact (Steinmuller, 2000).

Moving from different theoretical premises, Lafourcade and Paluzie (2011) show that border regions of core areas may obtain trade advantages from the integration process, as compared to other border regions since they can perform better in cross-country trade exchanges than interior regions. However, the issue of cross-country patterns of exchanges is important not only as far as the flows of goods are concerned. Knowledge flows, especially when tacit, can, as well, be characterized by differential patterns in border and in interior regions.

The so called French School of Proximity claims that geographical proximity is neither necessary nor sufficient for knowledge spillovers and that a separate role for a-spatial links among economic entities is possible (see Carrincazeaux and Coris, 2011). Such links have been classified by Boschma (2005) into five dimensions of proximity across agents: geographical, institutional, technological (or cognitive), social (or relational) and organizational. Several recent works have proved the relative importance of a-spatial dimensions on either economic performance (Basile et al., 2012) or on innovative activity (Marrocu et al., 2013).

In view of the arguments elaborated so far, we can now spell out our working hypotheses as it follows.

Hypothesis 1. Knowledge flows are affected by multi-faceted proximity. However, knowledge flows are not all alike and the diverse kinds of proximity have, consequently, differential impacts. Citations and co-inventorship may be thought as standing at two poles of a continuum marked by codified and tacit knowledge respectively. In this direction, co-inventorship is expected to be more sensitive to geographical proximity than citations, while the latter are expected to be more sensitive to technological proximity than the former.

Hypothesis 2. Being near an international border implies international contiguity which creates a better environment for knowledge exchanges with other regions in nearby countries. We therefore expect that inner regions, in countries which share a border with other countries, are less prone to exploit knowledge flows than border regions.

3 Data and Variables

3.1 The dataset

In order to obtain information on citation patterns, co-inventorship and applicant-inventor relationships, we use data extracted from the OECD REGPAT Database and the OECD Citations Database (January 2012). The former database presents patent data that have been linked to regions utilizing the addresses of the applicants and inventors. Two main dataset are covered by REGPAT: patent applications filed to the European Patent Office (EPO) and patent applications filed under the Patent Cooperation Treaty (PCT) at international phase. The OECD Citations database provides information on patent citations found in patent applications filed directly to the EPO or via the PCT. The geographical coverage relates to 276 NUTS2 regions located in 29 European countries (the EU-27 countries plus Norway and Switzerland)[1]. The reference period is the priority year: since it corresponds to the first filing worldwide and it is considered the closest date to the invention.

The REGPAT database is used in order to build the inter-regional matrices on co-inventorships and applicant-inventor links, while this database has to be combined with the Citations database in order to build the matrix on citation flows made and received by each region. Patent applications of citing and cited patents are, as a matter of fact, linked to regions on the basis of inventors' address thanks to the information provided in the REGPAT database. In case of multiple inventors, a proportional share is assigned to each region and, as a result, cells are not going to be made of integers.

It is important to emphasize that the majority of citations at EPO comes from patent examiners during their searches rather than from patent applicants and inventors (Criscuolo and Verspagen 2008). Nonetheless, since we aggregate citations to proxy knowledge interactions among regions rather than inventors' contacts, this issue becomes less crucial (Breschi and Lissoni 2006). In other words, we believe that, even though examiners play an essential role in the citation process at EPO, it is reasonable to assume that professionals in R&D laboratories know existing patents (that is public knowledge) in their fields.

As for collaborations in inventive activity we consider all those cases where patents have more than one inventor and they reside in different regions in Europe. For each patent, we first link each inventor's region to all the other regions of the same patent. To every pair of regions is then assigned a weight which is inversely proportional to the total number of pairs created for each patent. The final matrix is made of the sum of weights for all the regions pairs for all the patents considered.

Finally, as far as the relationship among applicants and inventors of the same patent is concerned, we consider those patent applications where at least one applicant and/or at least one inventor reside in different regions in Europe. In this case, patents are linked to regions by utilizing the addresses of the applicants and inventors. In case of multiple applicants and/or inventors, a proportional share is attributed. More detail on the construction of these two latter matrices can be found in Maggioni et al. (2011).

It is worth noting that while all of the three measures are different in that they capture different mehanisms underlying the dynamics of knowledge flows[2], the citation and the applicant-inventor matrices share a key property in that both kinds of flows are bi-directional, that is one can discriminate between the origin region and the destination region. On the contrary, with the co-inventorship matrix flows are not bi-directional and therefore there is no difference between origin and destination regions. This matrix is, in other words, symmetric. This implies that in the former two cases the number of observations is 75900 (n*(n-1): 276*275), whilst in the latter case we have to halve this number to obtain 37950 (n*(n-1)*1/2).

Table A1 shows the countries included in the analysis, as well as some key figures on patent activities and collaboration patterns.

3.2 The Variables

In this section we provide the definitions concerning the dependent and the explanatory variables used in the analysis. Details about the econometric methodology can be found in the Appendix A1.

3.2.1 Endogenous variables

In order to investigate the effects of the multifaceted dimensions of proximity on the different kinds of knowledge flows we use three different dependent variables: a) $Ln(coinv_{i,j})$ is the natural logarithm of co-inventorship collaboration between regions *i* and *j*, identified when, in a patent developed by more than one inventor, at least one co-inventor is resident in region *i* and at least one co-inventor is resident in region *j*; b) $Ln(appinv_{i,j})$ is the natural logarithm of applicant-inventor link between regions *i* and *j*, identified whenever a patent has (at least) one inventor in region *i* and one applicant (which is usually a firm) resident in another region *j*[3]; c) $Ln(cit_{i,j})$ is the natural logarithm citations link between regions *i* and *j*, which occurs when the citing patent has at least one inventor residing in the region *j* and the cited patent has at least one inventor residing in the region *i*.

As already emphasized, these three different indicators capture different knowledge flows dynamics. Actually, while patent citations can be interpreted as a proxy for knowledge externalities (insofar as the citing and the cited patent applicants are different firms/persons), the same does not apply to co-inventorship and applicant-inventor relationships, which often represent employer-employee relationships or colleagues working together in the same research team.

Moreover, the three indicators can be thought as mapping onto different kinds of knowledge defined according to the tacit/codified distinction. As it is shown in Figure 1, one can imagine tacit and codified knowledge as two separate poles of a continuum. In this frame, citation links better proxy the flow of codified knowledge between two regions, while co-inventorship is mostly related to the exchange of tacit knowledge. The link applicant-inventor can instead be seen as a sort of intermediate collaboration form. Actually, applicants are usually companies[4], and the kind of link established between an applicant and an inventor is much more similar to an employer-employee relationship than to collaboration. However, a successful innovation process requires not only skilled inventors, but also qualified employers able to screen and monitor inventors' activities. The sharing of some codified knowledge is therefore crucial. At the same time, the interactions between applicants and inventors are also sensitive to tacitness insofar as the invention leading to the patent application emerges as a specific and idiosyncratic outcome.

>>> INSERT FIGURE 1 ABOUT HERE <<<

Finally, the dependent variable is the log of average values in the period 2002-2004 of the three types of knowledge flows detailed above. All explanatory variables but the dummies are, on the contrary, calculated in a previous period, that is 1999-2001, in order to partially avoid potential endogeneity problems. By way of robustness check, we also run estimations with different lag specifications. We regress in particular the log of average values of knowledge flows in the period 2005-2007 against the average values of explanatory variables in the period 1999-2001 and against the average values of explanatory variables in the period

2002-2004. It is worth noting that our empirical setting implies the implementation of cross-section econometric estimations.

Moreover, we further test results' sensitivity to different specifications of the model with respect to three technological domains which may produce heterogeneous kinds of knowledge and corresponding flows. Two fields are chosen among Key Enabling Technologies (European Commission, 2012),: ICT and Biotechnology; whilst the third one refer to a high-tech sector with an established technological: Pharmaceuticals[5].

3.2.2 Explanatory variables

The variety of dimensions related to proximity have been measured by a number of indicators. First of all, geographical distance ($geodist_{i,j}$) is measured by logarithm of the row-normalized distance between regions i and j. Secondly, we build up a contiguity matrix ($cont_{ij}$) between regions i and j. We further decomposed the contiguity measure so as to appreciate the difference between contiguity of regions belonging to the same country ($wtbrd_{ij}$) and contiguity of regions belonging to different countries ($crossbrd_{ij}$) (see Figure A1 in the Appendix for a synthesis). Finally, we follow Lafourcade and Paluzie (2011) and analyze whether border regions (usually peripheral regions) are better off than inner regions (usually core regions) in exchanging knowledge with neighbour countries. To do so, we calculate one more dummy variable, i.e. $inner_{ij}$ which is equal to 1 if regions i and j are not contiguous but belong to two contiguous countries, and 0 otherwise (see Figure A1 in the appendix).

Table A2 (in the appendix) allows to grasp the magnitude of the former two distinct phenomena and to observe some very interesting facts. Applicant-inventor links (which are reported in the first four columns) in Germany, for example are mainly intra-national (83%), and consist of contiguous German regions for a significant quota (36%). Across-border links in

Germany are, therefore, only 17% in contrast with the opposite case of Ireland where we find the highest quota of international links, equal to 88%. When we consider citations, in the middle of the table, we find that Germany is a much more international player with a quota of intra-national citations of 56% and of international ones of 44%. Amongst the most innovative countries, the one which shows the highest propensity to cross-citations with other countries is Switzerland with a quota of 83% (of which only 5% between contiguous regions). Other very open countries are those ones with a negligible number of patents, such as Romania and Bulgaria. These countries appear to be rather internationalized also with respect to co-inventorships, with quotas of 100%. It is worth noting that Germany again shows mainly nationwide networks (85% of inventors cooperations are within borders), whilst Switzerland is quite open with an equal distribution of intra and inter-national co-inventorships. In an intermediate position we find other important innovative countries, such as Sweden and Finland with a quota of international co-inventorships of 40% and 32%, respectively.

As far as the other dimensions of proximity are concerned, we focus only on technological and institutional proximity and we dismiss organisational and social proximity for a twofold reasons. On the one hand, previous works (Maurseth and Verspagen, 200x and Paci and Usai, 2009,) have reported that the former two aspects are relatively important in determining flows of knowledge across regions. On the other hand, the latter two dimensions, social and organizational, are very difficult to capture at the regional level (Marrocu et al 2013).

$Techprox_{i,j}$ is the technological proximity between regions *i* and *j*. It draws upon Jaffe's cosine index (Jaffe, 1986 and 1989) and is based on the technological classes (technologies henceforth) to which patents are assigned[6] (see Appendix A2 for details).

Institutional proximity is usually measured in a much simpler way (see Marrocu et al, 2013): a dummy which is equal to unity when region *i* and *j* belong to the same country or zero viceversa. In other words, the sharing of the same legal framework and common culture is a proxy for institutional proximity. Such common background is bound to affect transaction costs and make knowledge exchange easier and less costly.

In line with gravity models, we also consider a number of phenomena which are meant to account for the masses of the two regions *i* and *j*. First of all, we include the population density (*dens*) of sampled regions, calculated as the ratio between the number of inhabitants and area (land use). Since we are focusing on knowledge flows, regions' attraction degree may depend on the local availability of human capital (*loghk*), which is the natural logarithm of people with tertiary education attainment. In the same vein we also include the natural logarithm of regional R&D expenditure (*logrdexp*) and of regional patent stock (*kcap*), which is the stock of patents calculated by applying the permanent inventory method to patent applications.

In Table A3 in appendix we provide a synthetic account of the variables used in the econometric estimations, as well as of the time period over which they have been calculated and the different data sources. We also report the descriptive statistics concerning both the endogenous and the explanatory variables.

Table A4 shows instead the Spearman correlation coefficients amongst the variables included in the empirical analysis.

4 Econometric results

In order to analyze the effects of the different dimensions of proximity, we have estimated a log-linear transformation of the gravity equation, as in equation (2). Firstly, we present the whole set of results for each type of knowledge flow estimation in table 1. We report a set of four models per type of knowledge flows which starts with a basic estimation of the gravity model with only geographical distance and the controls' set for regional characteristics.

The other models follow (in columns 2 to 4) with sequential complications of the explanatory set.

>>> INSERT TABLE 1 ABOUT HERE <<<

The first column shows the baseline model, and the geographical distance coefficient is negative and significant, as expected. Column (2) reports the results after the inclusion of technological (*techprox*) and institutional proximity (*instprox*), which have both the expected positive and significant coefficient. Note that this inclusion lowers the impact of both geographical distance and contiguity, which are nevertheless still significant. Column (3) reports the results of the estimations where contiguity is disentangled in international and intra-national bordering. Both indicators of adjacency have positive and significant coefficients, although the impact of within-border contiguity is higher than that of cross-border one. Finally, in column 4 we complete our estimation by adding the dummy *inner*, which is positive but not significant which implies that being at the border or not does not affect the patterns of cross country citation flows.

Results for applicant-inventor links are analogous: the coefficient on distance is negative and significant as expected (column (5)) and. Column (6) includes technological and institutional proximity, both showing a positive and significant coefficient. It is worth noting that in this case the coefficient of institutional proximity (*instprox*) is three times the one of technological proximity. Column (7) shows the estimations including the within-border and cross-border contiguity where, as expected, the coefficient of the former is far higher than that of the latter. In column (8) we include the variable *inner*, which is now negative and significant. Since applicant-inventor relationships involve more tacit exchanges than citations, being or not a border region matters: cross-border regions are better off in international knowledge flows than inner regions.

Finally, the last four columns show the results of estimations aiming at assessing the impact of proximities co-inventorships. Column (9) reports the baseline specification in which only geographical distance is taken into account. The coefficient is negative and significant, as expected.

In column (10) we introduce technological and institutional proximity which have positive and significant coefficients with the latter having a prevailing impact with respect to the former. In column (11) we dig into the differences between cross-border and within-border contiguity, by obtaining results consistent with the previous estimates, *i*.e. suggesting that the latter has a higher impact than the former. We finally include the dummy *inner* in column (12), which is also in this case negative and significant, suggesting that border regions are better off than inner regions when international co-inventorship links, with their tacit content, are at work.

4.1 The differential impact of proximity along time and across technological domains

The estimated impact of the proximity measures across the three measures of knowledge flows provides an interesting outline of the differences amongst tacit and codified knowledge exchanges. However, one can expect this impact to be time and technology specific. In order to check the sensitivity of our results to the specification of the period of observation and to the aggregation of different technologies, we run further econometric estimations that allow us to account for these aspects[7].

Tables 2 and 3 report standardised coefficients of our preferred model, i.e. columns (4), (8) and (12) of the table above, to assess the differential impact of proximities along time and across technological domains. Primarily, however, standardised coefficients allow a direct comparison of the impact of geographical distance, contiguity (in its three different qualifications), technological and institutional across knowledge flows across estimations with different dependent variables.

Let us, thus, focus on the first three columns of Table 2.

>>> INSERT TABLE 2 ABOUT HERE <<<

Results comply with our conceptual framework and therefore with our expectations. In particular, we find that in the first model the standardized coefficient of geographical distance is quite similar for the three types of flows, even though distance produces a higher negative impact on co-inventorship than on other knowledge flows. Results are more clearly differentiated when contiguity is considered: contiguity dummies show that the within-border and the cross-border contiguity yield more significant impacts on co-inventorship (column 3) than applicant-inventor links (column 2) and citation flows (column 1). A similar result is also found for the *inner* variable which has no significant impact on citations, while it has a higher impact on co-inventorship than on applicant-inventor links. All in all, these results suggest that the higher the tacit content of knowledge flows, the more sensitive they are to contiguity.

The situation is quite different with respect to technological proximity, which produces a lower impact on applicant-inventor links, an intermediate one on co-inventorship and a higher one on citation flows. The effect related to 'epistemic communities' makes therefore citation flows more sensitive to cognitive similarity.

Finally, institutional proximity yields a comparable impact on co-inventorship and applicant-inventor links and a much lower one on citation flows.

The results discussed above show clear-cut patterns as far as the differential impact of multidimensional proximity on diverse kinds of knowledge flows are concerned. An interesting issue concerns the robustness of this evidence to different lag specifications and a different time period. Columns (4) to (6) report standardized coefficients of the estimations of the determinants of knowledge flows of the period 2005-07 while keeping constant the reference period of the explanatory variables. The coefficients for geographical distance change only marginally: now applicant-inventor links clearly show the lowest coefficient, followed by citations and then co-inventorship. As far as contiguity is concerned, coefficients for co-inventorship are still higher than those for applicant-inventor links, and in turn than those of citations. Technological proximity is basically not affected, even though coefficients for co-inventorship and applicant-inventor links are now very similar. The same applies to institutional proximity.

The last three columns of Table 2 reports the results obtained by regressing knowledge flows over the period 2005-2007 against exogenous variables calculated over the period 2002-2004, i.e. by reproducing the same lag structure as the baseline estimations. Results are overall in line with those of the previous estimations, suggesting that both envisaged relationships between the variables and differences amongst the diverse knowledge flows are fairly robust to different temporal specifications.

Table 3 shows results for the baseline model for all fields and its replication for three technological domains: ICT, Biotechnology and Pharmaceuticals. Impacts are remarkably similar across domains. There are nevertheless some discrepancies which are worth noting. As for the impact of contiguity, this is clearly higher for co-inventorships in all domains but its influence is particularly high in ICT and Pharmaceuticals when measured across borders and only in ICT, when we relate to within borders flows. Technological proximity, on the contrary, produces analogous impacts across flows as in all domains, in ICT and Pharmaceuticals whilst the importance of cognitive proximity is remarkably similar across flows in Biotechnology. Finally, the influence of institutional proximity is clearly more important in ICT (across all types of flows) than in Biotechnology and Pharmaceuticals.

In conclusion these last results show that while the model is certainly robust, knowledge flows are heterogeneous with respect not only to the channel used to move across firms, region and countries but also to the technological field at stake.

INSERT TABLE 3 ABOUT HERE

5 Conclusions and Policy Implications

Knowledge flows are not all alike. This is the answer to our main question, based on an empirical test which has assessed the functioning of three types of knowledge flows: citations, applicant-inventor links and co-inventorships. This is even more evident when we look at the role of proximity on the different indicators across diverse technological fields. More specifically, the estimation of a set of gravity models show that knowledge flows are affected by contiguity and proximities to different extents. We prove that, depending on the content of tacitness entailed in the knowledge flow, physical distance and more precisely contiguity may play a very different role. The highest impact of contiguity (both within and across countries) is registered for co-inventorship collaborations, that is those flows which are essentially based on tacit knowledge, cooperation and trust and are facilitated by face to face contacts. Consequently, facial contacts, and therefore contiguity, are less important for applicant-inventors links and are the least important for citations flows, since they are less dependent on personal contacts.

Sharing the same institutional context has also a diverse impact on knowledge flows as it is more important for collaborations among inventors and for applicant-inventors relationships whilst it is relatively less important for citations links. The rationale for this result is that in the former two cases a common institutional framework reduces the uncertainty and makes exchanges among economic agents less risky. The effects of contiguity and of institutional closeness are associated when we discriminate between contiguous regions within the same country and those which share an international border. As expected, being contiguous within national borders implies a stronger impact with respect to the case of contiguity across borders.

Knowledge flows which happen thanks to citation links are, on the contrary, more influenced by technological relatedness than the other two knowledge flows. This confirms that some elements of knowledge flow more easily within epistemic communities which share codified knowledge thanks to some rules for knowledge diffusion and they convey messages to whatever distance and independently from contiguity (see Breschi and Lissoni, 2001).

Finally, international border regions are shown to have an advantage with respect to other regions within the same country which are not on the border. This implies that bordering regions can emerge as more central thanks to their cross-border nature, and this effect may counteract at least partially the diseconomies due to peripherality.

The results of this study bear important implications for future research avenues and for technology policy. Actually, our findings are consistent with the evidence put forth by Breschi and Lissoni (2009), according to which the localization of knowledge spillovers is mainly driven by the limited mobility of inventors and their embeddedness in spatially bounded knowledge networks. Regional technology policies aiming at stimulating interregional or international knowledge spillovers should therefore be combined with migration and labour market policies able to *remove obstacles and frictions to the mobility of skilled labour force*, or to attract excellent researchers in the area. Also, cultural differences and institutional proximity affect more co-inventorship and applicant-inventor relationship than citation flows. This is probably due to the face-to-face nature of the first two types of knowledge exchanges. A possible avenue to mitigate this issue can rest on the *promotion of cultural diversity* in local contexts, extending the concept of absorptive capacity so as to include the capacity to manage interactions with people from different nationalities next to the ability to understand and internalize knowledge produced elsewhere. Moreover, the differences in the impact of geographical proximity on the three indicators seem to narrow over time, while this does not apply to technological and institutional proximity, the effects of which are persistent over time in terms of magnitude across the three measures. This suggests that, also due to increased speed and lower costs of transport, physical closeness becomes less and less important for co-inventorship, even though it remains significant. The real difference resides in the need for co-inventors to be able to communicate to exchange knowledge, i.e. the sharing of common codebooks, complementary or similar skills and competences. Once again, the key issue is the capacity to absorb tacit knowledge.

The evidence that knowledge flows are not all alike also opens up further research avenues. First of all, the impact of cultural diversity on knowledge flows needs to be more directly scrutinized, so as to validate our interpretation of the findings of the present study and to shed a new light on the economic analysis of migration patterns from the viewpoint of economics innovation. Moreover, the logical follow up of this analysis should concern the differential impact of the diverse types of knowledge flows on regional innovation performances and then on regional differentials in productivity growth, so as to ascertain the extent to which regional technology policies need to promote the exchange of codified rather than tacit knowledge to achieve higher levels of per-capita income.

References

Allen, R.C., (1983). "Collective invention," *Journal of Economic Behavior & Organization*, 4(1), 1-24,

Anderson J. E. and van Wincoop E., (2004) Trade Costs, *Journal of Economic Literature*, 42, 691–751.

Antonelli, C. (1999). *The microdynamics of technological change*. London: Routledge.

Arrow K.J., (1962) The Economic Implications of Learning by Doing, *Review of Economic Studies*, 29, pp. 155-173.

Audretsch D.B. and Feldman M.P., R&D spillovers and the geography of innovation and production, *American Economic Review*, (1996) 86, 630–640.

Baptista, R. and Swann, P., (1998). "Do firms in clusters innovate more?," *Research Policy*, 27(5), 525-540,

Basile R. and Usai S., (2012). "Analysis of regional endogenous growth," In Karlsson C., Andersson M. and Norman T. (eds): Handbook of Research Methods and Applications in

Economic Geography. Edward Elgar, Cheltenham Belenzon S. and Schankerman M. (2013), Spreading the word: Geography, Policy and Knowledge Spillovers, *The Review of Economics and Statistics*, July 2013, 95(3): 884–903

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

Breschi, S. and Lissoni, F. (2001) Knowledge Spillovers and Local Innovation Systems: A Critical Survey, *Industrial and Corporate Change*, 10, 975-1005.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association* 83, 123–127.

Cappelli, R. and Montobbio, F., (2016). "European Integration and Knowledge Flows across European Regions," Regional Studies, 50, 709-727.

Carrincazeaux C. and Coris M., (2011) Proximity and Innovation, in Cooke P., Asheim B.T. and Boschma R. (eds) *Handbook of Regional Innovation and Growth*. Cheltenham: Edward Elgar.

Costantini, V. and Crespi, F., (2008). "Environmental regulation and the export dynamics of energy technologies," *Ecological Economics*, 66(2-3), 447-460,

Cowan, R., David P. A. and Foray, D., (2000). "The Explicit Economics of Knowledge Codification and Tacitness," *Industrial and Corporate Change*, 9(2), 211-53.

Criscuolo, P. and Verspagen, B., 2008. "Does it matter where patent citations come from? Inventor vs. examiner citations in European patents," *Research Policy*, 37(10), 1892-1908.

Engelsman, E. C., and Van Raan, A. F. J. (1991). *Mapping Technology. A first exploration of knowledge diffusion amongst fields of technology*. The Hague: Ministry of Economic Affairs.

European Commission (2012), COM(2012)-341, Final Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee Of The Regions 'A European strategy for Key Enabling Technologies –A bridge to growth and jobs'.

Feldman, M.P. (1994a). *The Geography of Innovation,* Kluwer Academic Publishers, Boston.

Feldman, M.P. (1994b), "Knowledge complementarity and innovation, Small Business Economics 6(3): 363-372.

Foray, D. (2004). *The economics of knowledge*. Cambridge, Mass.: MIT Press

Jaffe A.B., Real Effects of Academic Research, *American Economic Review*, (1989) 79, 957-70.

Jaffe A.B. (1986) Technological Opportunity and Spillovers of R&D: evidence from Firms' Patents, Profits and Market Value, *American Economic Review*, 76, 984-1001.

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

Jaffe, A.B, Fogarty M.S and Banks B.A., (1998) "Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation*," Journal of Industrial Economics*, 46(2), 183-205.

Johnson, N.L. (1949). Systems of Frequency Curves Generated by Methods of Translation. *Biometrika* 36, 149-176.

Lafourcade M. and Paluzie E., (2011). European Integration, Foreign Direct Investment (FDI), and the Geography of French Trade, *Regional Studies*, 45(4), 419-439.

Lundvall, B-Å. (ed.) (1992), *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*, London: Pinter Publishers.

Maggioni M.A., Nosvelli, M. and Uberti, T.E. (2007) Space versus networks in the geography of innovation: a European analysis, *Papers in Regional Science*, 86, 471-493.

Maggioni M.A., Uberti T.E. and Usai S. (2011) Treating patents as relational data: knowledge transfers and spillovers across Italian provinces, *Industry & Innovation*, 18, 39-67.

Maraut S., Dernis H., Webb C., Spiezia V and Guellec D., (2008). "The OECD REGPAT Database: A Presentation," *OECD STI Working Papers* 2008/2

Marrocu E., Paci R. and Usai S. (2013) Proximity, Networking and Knowledge Production in Europe: what lessons for innovation policy? *Technological Forecasting and Social Change,* 80 (2013), pp. 1484-1498,

Maurseth, P. B. and Verspagen, B., (2002). " Knowledge Spillovers in Europe: A Patent Citations Analysis," *Scandinavian Journal of Economics*, 104(4), 531-45,.

Meade J., (1952), *A Geometry of International Trade*, London, George Allen & Unwin, 1952.

Montobbio F. and Sterzi V., (2013) The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations, *World Development*, 44, 281-299.

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

OECD (2009), *OECD Patent Statistics Manual*, OECD Paris

Paci R. and Usai S. (2009) Knowledge flows across the European regions, Annals of Regional Science, 43, 669-690.

Paci R., Marrocu E. and Usai S. (2014) The complementary effects of proximity dimensions on knowledge spillovers, *Spatial Economic Analysis*, 9, 9-30.

Picci L. (2010) The internationalization of inventive activity: A gravity model using patent data, *Research Policy*, 39, 1070-1081.

Santos Silva J.M.C. and Tenreyro S., (2006), The Log of Gravity, *The Review of Economics and Statistics*, 88(4): 641–658

Saviotti, P.P, (1998). "On the dynamics of appropriability, of tacit and of codified knowledge," *Research Policy*, 26(7-8), 843-856.

Sorenson O. and Singh J., (2007). "Science, Social Networks and Spillovers," *Industry and Innovation*, 14(2), 219-238.

Steinmueller, W E., (2000). "Will New Information and Communication Technologies Improve the 'Codification' of Knowledge?," *Industrial and Corporate Change*, 9(2), 361-76.

Strumsky, D., Lobo, J., Van der Leeuw, S. (2011) "Using Patent Technology Codes to Study Technological Change" *Economics of Innovation and New Technology*, 1-20

Thompson P., (2006). "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations," *The Review of Economics and Statistics*, 88(2), 383-388,

Thompson P. and Fox-Kean M., (2005). "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *American Economic Review* 95(1), 450-460,

Viner, J. (1931). Costs curves and supply curve. *Zeitschrift für Nationalokonomie* 3: 23-46.

Von Hippel E. (1988), *The Sources of Innovation*, Oxford University Press

Von Hippel E. (1994) "Sticky Information" and the Locus of Problem Solving: Implications for Innovation, *Management Science*, 40, 429-439.

Table 1 – Econometric Results.

	Citations				AppInv				Coinventorship			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
geodist	-0.1342***	-0.0697***	-0.0706***	-0.0714***	-0.2210***	-0.0526***	-0.0572***	-0.0617***	-0.1832***	-0.0450***	-0.0484***	-0.0687***
	(0.0037)	(0.0037)	(0.0037)	(0.0042)	(0.0056)	(0.0039)	(0.0038)	(0.0041)	(0.0094)	(0.0056)	(0.0055)	(0.0063)
contig		0.2073***				0.5450***				0.5205***		
		(0.0324)				(0.0441)				(0.0660)		
techprox		0.2543***	0.2543***	0.2543***		0.1559***	0.1557***	0.1556***		0.1617***	0.1636***	0.1665***
		(0.0226)	(0.0226)	(0.0226)		(0.0203)	(0.0201)	(0.0201)		(0.0308)	(0.0306)	(0.0305)
instprox		0.1999***	0.1943***	0.1927***		0.5275***	0.5003***	0.4917***		0.4683***	0.4402***	0.3919***
		(0.0142)	(0.0146)	(0.0143)		(0.0166)	(0.0169)	(0.0166)		(0.0249)	(0.0249)	(0.0255)
crsbrd			0.1473***	0.1451***			0.2557***	0.2433***			0.3131***	0.2558***
			(0.0449)	(0.0450)			(0.0505)	(0.0504)			(0.0809)	(0.0791)
wtnbrd			0.2288***	0.2280***			0.6489***	0.6446***			0.6178***	0.6047***
			(0.0406)	(0.0406)			(0.0560)	(0.0559)			(0.0880)	(0.0871)
inner				-0.0026				-0.0144***				-0.0622***
				(0.0065)				(0.0053)				(0.0083)
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	75900	74256	74256	74256	75900	74256	74256	74256	37950	37128	37128	37128
R^2	0.354	0.376	0.376	0.376	0.245	0.345	0.348	0.348	0.317	0.451	0.454	0.457
adj. R^2	0.354	0.375	0.375	0.375	0.245	0.345	0.347	0.347	0.314	0.448	0.451	0.454
AIC	48874.9823	46714.4311	46703.8159	46705.4452	61909.3957	51542.0500	51268.4329	51259.5762	3163.0063	406.4640	330.9911	264.3810
BIC	49484.6356	47350.2850	47348.8851	47359.7296	62519.0490	52177.9039	51913.5021	51913.8606	3450.2143	700.2088	632.2678	580.7215

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2 – Standardized Coefficients. Comparison amongst different lag specifications.

	Dep. Var.: 2002-2004 – Expl. vars: 1999-2001			Dep.Var.: 2005-2007 Expl. vars: 1999-2001			Dep. Var.: 2005-2007 – Expl. vars: 2002-2004		
	Ln(Cit)	Ln(AppInv)	Ln(Coinv)	Ln(Cit)	Ln(AppInv)	Ln(Coinv)	Ln(Cit)	Ln(AppInv)	Ln(Coinv)
geodist	-0.120***	-0.103***	-0.146***	-0.115***	-0.076***	-0.103***	-0.113***	-0.073***	-0.102***
techprox	0.052***	0.032***	0.043***	0.050***	0.036***	0.034***	0.050***	0.037***	0.034***
instprox	0.115***	0.292***	0.289***	0.123***	0.278***	0.306***	0.124***	0.279***	0.306***
crsbrd	0.022***	0.036***	0.058***	0.029***	0.049***	0.064***	0.030***	0.049***	0.064***
wtnbrd	0.061***	0.170***	0.218***	0.069***	0.197***	0.292***	0.070***	0.197***	0.292***
inner	-0.002	-0.013***	-0.076***	0.001	-0.020***	-0.074***	0.002	-0.019***	-0.074***
N	74256	74256	37128	74256	74256	37128	74256	74256	37128

Standardized beta coefficients;
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3 – Standardized Coefficients. Comparison amongst different sectors

	All fields			Biotechnology			ICT			Pharmaceuticals		
	Cit	AppInv	Coinv	Cit	AppInv	Coinv	Cit	AppInv	Coinv	Cit	AppInv	Coinv
geodist	-0.120***	-0.103***	-0.146***	-0.022***	-0.035***	-0.078***	-0.052***	-0.045***	-0.076***	-0.037***	-0.042***	-0.084***
	(-17.03)	(-14.95)	(-10.84)	(-3.25)	(-3.99)	(-6.37)	(-6.84)	(-5.42)	(-7.73)	(-5.53)	(-5.11)	(-7.35)
techprox	0.052***	0.032***	0.043***	0.020***	0.021***	0.023***	0.051***	0.031***	0.039***	0.033***	0.019***	0.024***
	(11.24)	(7.73)	(5.46)	(4.58)	(4.86)	(3.67)	(9.56)	(6.92)	(6.18)	(6.56)	(4.69)	(4.07)
instprox	0.115***	0.292***	0.289***	0.056***	0.134***	0.148***	0.086***	0.180***	0.209***	0.046***	0.123***	0.150***
	(13.43)	(29.54)	(15.39)	(5.44)	(10.45)	(9.84)	(8.44)	(14.65)	(14.51)	(5.04)	(10.64)	(10.14)
crsbrd	0.022***	0.036***	0.058***	0.008	0.018*	0.038*	0.009	0.026***	0.054***	0.016*	0.024*	0.059**
	(3.23)	(4.83)	(3.23)	(1.49)	(1.91)	(1.90)	(1.57)	(3.64)	(3.83)	(1.73)	(1.95)	(2.19)
wtnbrd	0.061***	0.170***	0.218***	0.034***	0.124***	0.196***	0.059***	0.161***	0.280***	0.058***	0.126***	0.195***
	(5.62)	(11.53)	(6.95)	(2.77)	(5.85)	(5.77)	(4.44)	(8.65)	(9.08)	(3.87)	(6.42)	(6.28)
inner	-0.002	-0.013***	-0.076***	0.002	-0.018***	-0.042***	-0.008	-0.020***	-0.058***	-0.007	-0.011**	-0.036***
	(-0.40)	(-2.71)	(-7.51)	(0.40)	(-3.44)	(-5.36)	(-1.30)	(-3.99)	(-8.34)	(-1.23)	(-1.99)	(-4.93)
N	74256	74256	37128	74256	74256	37128	74256	74256	37128	74256	74256	37128

Standardized beta coefficients;
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[1] Data on patents in the OECD REGPAT database (Maraut et al., 2008), provides information on inventive activity and its multiple dimensions (e.g. geographical location, technical and institutional origin, individuals and networks).

[2] We emphasize that knowledge flows not necessarily imply spillovers: co-inventorships and applicant-inventors links operate most of the time within the boundaries of a firm, which therefore internalizes potential externalities.

[3] In some patents the applicant can be the inventor him/herself. This does not create any problem in this context, as in these cases inventor and applicant appear to belong to the same region and therefore they are not counted.

4 The case in which the applicant id is the same as the inventor id is not taken into account by definition.

[5] These domains are identified by OECD (2009) for analytical and political interests, on the basis of the available information: the IPC code and/or the textual data.

[6] See Strumsky et al., 2012, for a critical assessment of opportunities and shortcomings related to the use of technological classes in empirical analyses.

[7] By way of robustness check we also run additional estimations to control for the impact of i) regions' size by considering the GDP levels instead of Employment or R&D expenditure; ii) regional dummies together with pair-level mass variables instead of regional variables; iii) the impact of contemporaneous knowledge stock instead of the lagged one; iv) the dummy for the sharing same languages. The results are reported in the Appendix, Table A5. While the estimations yield significant coefficients with the expected sign for these variables, the key results on the proximity measures mostly persist and are consistent with those already discussed in Section 4.