Networks of innovators within and across borders. Evidence from patent data

Andrea Morescalchi
Fabio Pammolli
Orion Penner
Alexander M. Petersen
Massimo Riccaboni

Networks of innovators within and across borders. Evidence from patent data

Andrea Morescalchi
IMT Institute for Advanced Studies Lucca

Fabio Pammolli
IMT Institute for Advanced Studies Lucca

Orion Penner
IMT Institute for Advanced Studies Lucca

Alexander M. Petersen
IMT Institute for Advanced Studies Lucca

Massimo Riccaboni
IMT Institute for Advanced Studies Lucca and Department of Managerial Economics, Strategy and Innovation, K.U. Leuven

Networks of innovators within and across borders. Evidence from patent data

Andrea Morescalchi*[1], Fabio Pammolli[1], Orion Penner[1],
Alex M. Petersen[1], and Massimo Riccaboni[1,2]

[1]IMT Institute for Advanced Studies Lucca, 55100 Lucca, Italy
[2]DMSI, K.U. Leuven, Leuven, Belgium

July 22, 2013

Abstract

Recent studies on the geography of knowledge networks have documented a negative impact of physical distance and institutional borders upon research and development (R&D) collaborations. Though it is widely recognized that geographic constraints hamper the diffusion of knowledge, less attention has been devoted to the temporal evolution of these constraints. In this study we use data on patents filed with the European Patent Office (EPO) for 50 countries to analyze the impact of physical distance and country borders on inter-regional links in four different networks over the period 1988-2009: (1) co-inventorship, (2) patent citations, (3) inventor mobility and (4) the location of R&D laboratories. We find the constraint imposed by country borders and distance decreased until mid-1990s then started to grow, particularly for distance. The intensity of European cross-country inventor collaborations increased at a higher pace than their non-European counterparts until 2004, with no significant relative progress afterwards. Moreover, when analyzing networks of geographical mobility, multinational R&D activities and patent citations we do not depict any substantial progress in European research integration aside from the influence of common global trends.

---

*Corresponding author (andrea.morescalchi@imtlucca.it; andrea.morescalchi@gmail.com)

1 Introduction

Rapid progress in information, communication, and transportation technologies and the overall trend of globalization have lead to the assertion "distance is dead" (Castells, 1996; Cairncross, 1997). A natural tension exists between this view and knowledge "stickiness": human activities and social interactions are known to geographically cluster to take advantage of knowledge spillovers, social capital and other agglomeration economies (Feldman, 1994). While the literature on innovation systems has focused on the interplay between clusters and networks of innovators (Breschi & Malerba, 2005), the "death of distance" conjecture has been thoroughly investigated in the literature on international trade and globalization studies. The most significant recent advances in that vein have been made by means of panel gravity regressions and indicate distance, borders and free trade areas still play a key role in trade networks (the so-called "tyranny of distance"). Since the seminal contribution of Freeman (1991), networks of innovators have attracted a great deal of interest as a tool for representing and analyzing division of innovative labor. Many types of network have been investigated, ranging from the informal scientific connections in invisible colleges and communities of practice to the formal collaborative agreements between firms and other research organizations. With increasing frequency, growing data on scientific collaborations, collaborative R&D projects, and patents have been widely exploited to gain insight into the structure and evolution of networks in different industries, countries and timeframes (see Powell & Grodal, 2005 and Ozman, 2009 for reviews).

Despite significant efforts in a growing body of literature analyzing networks of innovators, there is still a lack of large-scale quantitative understanding of the role of geographic borders and distance. The complexity of this problem arises from the variety of competing forces that underlie the economics and sociology of R&D collaboration. Prevailing wisdom states the spread of tacit knowledge and the formation of informal ties are uninhibited over short distances, but barriers increase with distance. However, in the case of formal contractual collaborations and transmission of codified knowledge, distance plays less of a role at large scales, even if borders between different institutional settings can still reduce the effectiveness of contractual solutions. Because technological advancements have increased the capacity to codify and share knowledge across large distances, it follows that the barriers induced by distance should be decreasing, and possibly vanishing, in R&D networks. Also, the dynamic role of physical distance and institutional borders may differ significantly across different R&D networks depending upon the type of knowledge that is exchanged (tacit vs. codified) and the nature of the links: market transactions, hierarchical relations or network forms of coordination (Whittington et al., 2009). Moreover cross-network interdependencies should be taken into account. For example, international mobility should have a positive impact on regional citation flows as inventor movement is thought to be an important driver of knowledge flows. International mobility may, in turn, have a positive impact on large distance collaborations as mobile inventors act as bridges across teams of inventors working for different organizations (Breschi & Lissoni, 2009). Conversely, one could argue that, the more individual inventors and research teams can freely move, the less R&D organizations will feel the need to locate R&D labs abroad or to sign collaborative agreements with foreign partners. In this sense, understanding the extent to which globalization reduces constraints on geographical mobility is important for assessing side-effects in other dimensions of R&D networks.

Here we employ a gravity approach to quantify the strength of borders and distance on multiple innovation networks. We analyze a large sample of developed nations over many years to investigate the dichotomy arising from localizing constraints of R&D spillovers and agglomeration economies in R&D clusters vis-à-vis the tendency to expand R&D networks via long-range collaborations between inventors located in different countries and institutional settings.

Beyond scientific relevance, a better understanding of how distance and borders influence the

structure and evolution of R&D networks is important to orient the policy debate. In particular, the European Research Area (ERA) vision of an "open space for knowledge and growth" stands as the most recent in a long line of integration efforts within the European Union (EU). The ERA realization has been highlighted as key component of the competitiveness of the EU's Europe 2020 growth strategy. This is an attempt to reduce, perhaps even eliminate, the effect of national borders on scientific and R&D networks to create an area in which ideas and high skill human capital are free to flow and capitalize on transnational synergies and complementarities. Related to the creation and consolidation of an ERA, is the implementation of Research and Innovation Strategies for Smart Specialization (RIS3 strategies), a set of guidelines the European Commission has identified as a key element to promote integration of national and regional innovation efforts for the achievement of smart, sustainable and inclusive growth (see Foray & Van Ark, 2007 for a discussion of the concept of smart specialization).[1] The idea of smart specialization is based on the Marshallian notion that regions with production structures specialized towards a particular industry tend to be more innovative in that particular industry (Glaeser et al., 1992). This tendency follows because regional specialization promotes economies of scale, agglomeration, and spillovers in knowledge production and use, which are important drivers of productivity. Furthermore, a better understanding of the role distance and borders play in the structure and evolution of networks of innovators is key to not only for crafting effective policy, but even more simply, for assessing the true effectiveness of past, present, and future policy measures.

Our study moves beyond previous efforts to understand the geography of research collaboration in three key ways. First, we study a more comprehensive set of countries (50 OECD and OECD-partner countries) at a lower level of spatial aggregation (NUTS3).[2] Most previous studies used NUTS2 and the few that used NUTS3 focused on a single country (Ponds et al., 2007; Frenken et al., 2009b) or a few countries (an exception being Hoekman et al., 2009 who analyze EU27 countries plus Norway and Switzerland). Second, we study a set of interrelated patent networks using the same analytic approach: (1) the network of patent co-inventorship, (2) the location of R&D labs (the applicant-inventor network), (3) patent citations and (4) inventor mobility. This is in contrast to previous studies that generally focused only on one network at a time. Third, in our analysis we investigate, jointly, the distance effect (Ponds et al., 2007) and the country-border effect (Ponds, 2009) in Europe and other OECD countries. Few previous studies have investigated the dynamics of the distance and border-effects simultaneously (Hoekman et al., 2010; Singh & Marx, 2012) and all focus on either Europe or the United States. Our more comprehensive analysis allows us to examine and ultimately quantify the effect of European integration efforts, by applying a suitable counterfactual approach (Chessa et al., 2013).

This paper proceeds in the following manner. Section 2 presents a review of the relevant literature. In section 3 we describe the data and methodology used. Section 4 presents the results of our analysis. Finally, in section 5 we discuss our results and natural extensions of this research direction deriving some policy implications for the European Research Area.

---

[1] The relevance of RIS3 strategies is pointed out in the European Commission's proposal for cohesion policy in 2014-2020 (`http://ec.europa.eu/regional_policy/what/future/proposals_2014_2020_en.cfm`) in which their implementation is proposed as precondition for using the European Regional Development Fund (ERDF) in 2014-2020 to support research and innovation investments.

[2] The Nomenclature of Units for Territorial Statistics (NUTS) is a geo-code standard for referencing the subdivisions of countries for statistical purposes. The nomenclature has been introduced by the EU for its member states. The OECD provides an extended version of NUTS3 for its nonEU member and partner states.

2 The role of geography in networks of innovators

The prevailing wisdom is that globalization and advances in information, transportation, and communication technologies should reduce the role of distance in socio-economic interactions (Castells, 1996; Cairncross, 1997). This issue has been thoroughly explored in the literature on trade through the lens of gravity models (Coe, 2002; Brun et al., 2005). Contrary to predictions, results obtained from a wide variety of approaches and data have led to an an emerging consensus that distance still plays an important role in constraining trade flows. Recent studies of the geography of R&D networks have also documented the relevance and persistence of spatial biases.

Most previous analysis of the globalization of the knowledge production have focused on two specific spatial biases. First, the degree to which travel and communication costs result in physical distance being an impediment to collaboration. Second, the extent to which institutional friction arising from country-to-country differences create challenges for collaboration across national systems of innovation (Freeman, 1995; Lundvall, 1992; Nelson, 1993). Application of gravity models to scientific and technological collaboration have provided strong evidence for a negative effect of physical distance and country borders (Ponds et al., 2007; Maggioni & Uberti, 2007; Scherngell & Barber, 2009; Frenken et al., 2009b; Hoekman et al., 2009, 2010; Scherngell & Barber, 2011; Scherngell & Hu, 2011; Hoekman et al., 2013; Scherngell & Lata, 2012). This body of evidence is robust over various kinds of data (scientific publications, patents), the type of network (collaborations between individuals/institutions, citations, labour mobility) and the geographic unit of analysis (country, regional, sub-regional).

Given that costs of coordinating R&D activities at distance are rapidly decreasing and the general globalization of the science system, it is widely assumed that the bias to collaborate domestically, and the bias to collaborate at a close distance have been waning. This may be especially true for European countries, where specific steps have been taken at the political level to stimulate integration in R&D. Several studies have found evidence in favor of this assumption, but some scholars also found different results (see Frenken et al., 2009a for a survey). The conclusion that spatial biases are attenuating is generally arrived at after observing an increase in the cross-border shares of collaborations and an increase in the average distance of collaborations. However, using different methodologies, some studies provide evidence that the constrain of distance is becoming more binding over time (Hoekman et al., 2010; Singh & Marx, 2012; Ponds et al., 2007; Boerner et al., 2006), whereas others (Singh & Marx, 2012; Ponds, 2009; Frenken, 2002) show that the country-border effect is not lessening. Agrawal & Goldfarb (2008) studied the effect of a decrease in the cost of collaboration between university based engineering groups resulting from adoption of Bitnet (an early version of Internet). They found that in some sub-samples the greatest benefit was experienced by university pairs that were geographically close. One may expect that long distances collaborations would benefit more from improvement in communication technologies since their cost decreases the most, but this finding supports the view reduced communication costs can, indeed, accentuate tendencies for research activity to agglomerate rather disperse. Related to this evidence, Gaspar & Glaeser (1998) point out that telecommunications are not necessarily a substitute for face-to-face interactions. When telecommunications technology improves, we can expect that some interactions otherwise conducted face-to-face will instead be conducted electronically. However, it is also possible that such improvements result in an increased frequency of contact between individuals, necessitating further close interactions. It follows that if the second effect is sufficiently large, it may even be possible to observe an increase in the importance of spatial proximity as information technology improves.

Among all studies examining the dynamics of spatial biases, two in particular have employed a sound statistical approach. Hoekman et al. (2010) estimate gravity models using data on co-

publications between NUTS2 regions in 33 European countries for the period 2000-2007. They find that the negative effect of distance on inter-regional collaborations increases over the focus period and that the country-border effect decrease, though not statistically significantly.[3] Singh & Marx (2012) analyze citations to US patents applied for over the period 1975-2004, with the United States Patent and Trademark Office (USPTO). Using an approach in which the unit of analysis are pairs of patents representing actual and potential citations, and the probability of observing an actual citation is modeled with a weighted logistic regression, they observe an increase over time in the citations received from non-US patents relative to citations received from US patents.[4] That study also finds that the rate of decay in the probability of citation as a function of distance has slightly increased over time. That is the effect of distance is increasing.

While Singh & Marx (2012) finds evidence that, for the US, the role of national borders is increasing, Hoekman et al. (2010) find that in Europe the same effect has decreased over time. These are, indeed, opposite trends but were obtained considering different regions (US, EU respectively) and networks. Here we attempt to bring coherence to the issue of the dynamics of distance and borders by considering a broad range of countries and many different networks each with their own dependance upon tacit versus codified knowledge. Moreover we aim to uncover the possible effect of EU integration policies through a regression approach capable of determining evolution of the country-border effect for European versus non European countries. In particular, following Chessa et al. (2013), we readapt our methodological framework to assess whether policies oriented to promote cross-border collaboration in Europe have lessened the national border effect. Specifically we seek to test if (i) the effect of physical distance and (ii) country borders are decreasing, and (iii) if the country border effect is decreasing in Europe relative to the rest of the developed countries.

The patent networks we analyze are important representations of knowledge geography and provide quantitative structures for measuring knowledge diffusion. Since the pioneering work of Jaffe et al. (1993), patent citations have been utilized extensively to measure the diffusion of knowledge across a variety of dimensions: geographic space, time, technological fields, organizational boundaries, alliance partnerships, and social networks (Almeida & Kogut, 1999; Jaffe & Trajtenberg, 2002; Peri, 2005; Gomes-Casseres et al., 2006). A principal assumption underlying this approach is that citations trace out knowledge flows and technological learning as knowledge embedded in the cited patent is transmitted to inventors of the citing patent. Given that access to codified knowledge typically do not require interactions between individuals, it is recognized that distance and institutional borders should be relatively less important in this network. Such studies focus on citations as means to transfer codified knowledge but acknowledge that citations are less effective means of spreading tacit knowledge than personal, face-to-face contacts.

Though many empirical studies have analyzed the role of patent citations as measure of knowledge flows it has also been stressed that economic agents can access knowledge from many other sources than only codified knowledge. In particular, a distinction between two means of spreading tacit knowledge has been made in the literature, which operate either through informal social interactions, arm-length market-based relationships, inter-organizational alliances or hierarchical solutions within R&D organizations. Examples of the first case are social ties with current and former colleagues and those developed in social events (conferences, affiliation to associations etc.). Geography is relevant here as proximity facilitates the development of social relationships and raises incentives to invest in social capital (Agrawal et al., 2006). In the second case, the transmission of knowledge is regulated by a contract, such as a labour contract, licensing or formal collaborations, which explicitly set a compensation for the exchange of knowledge (Breschi & Lissoni, 2009). Geography matters either

---

[3]However they find a decreasing border effect as regards regional borders.
[4]They also find that the state-border effect decreases over time as consistent with Hoekman et al. (2010).

because labor mobility among different institutions or laboratories can be constrained in space, or because formal agreements require frequent interactions and monitoring that are more easily conducted locally. The network of co-inventions stands somehow in between these two categories as either the collaboration can be ruled by a formal agreement or inventors can decide to collaborate informally with colleagues located in different areas. The co-inventor network is affected by geography as spatial proximity and co-location may facilitate the transfer of complex knowledge as frequent face-to-face interactions maybe required. Though easing of communication and travel constraints is expected to reduce the importance of spatial proximity in this network, the result can depend on the degree of complementarity between remote and face-to-face interactions.

The popularity of patent citations and collaborations as a means to capture knowledge flows is probably motivated by the interest in economics for pure externalities (spillovers), *i.e.* a transfer of knowledge which is not mediated by the market (Breschi & Lissoni, 2009). The other two networks we present, *i.e.* relationships between organizations and affiliated inventors, and inventor moving across organizations or across regional laboratories within the same institution, operate through market-based channels.[5]

The geographical links between applicants and affiliated inventors is relevant to the analysis of the geographic distribution and globalization of the innovative activities of firms (see Keller, 2004 and Narula & Zanfei, 2005 for surveys). Multinational firms are well known to be drivers of the internationalization of innovation activities (see Wolfmayr et al., 2013) as international location of a firm's subsidiaries facilitates knowledge transfer across borders. The literature on the internationalization of business suggests a number of different reasons for undertaking technological activities outside the home country (Dunning & Lundan, 2009; Florida, 1997; von Zedtwitz & Gassmann, 2002). Among these, knowledge-seeking motives such as proximity to university and innovative firms as a means to benefit from spillovers and agglomeration advantages, and access to high quality scientific and technical talent, have become considered extremely relevant since the late 1990s (Florida, 1997; von Zedtwitz & Gassmann, 2002; Patel & Vega, 1999; Granstrand, 1999). For example, von Zedtwitz & Gassmann (2002) show that these knowledge related factors are by far the most important motives for performing "research" (rather than "development") activities at foreign locations. Indeed, localized foreign knowledge that is tacit can be accessed or imported for firms by moving closer to the source. This goal can be achieved by setting up subsidiaries abroad (Phene & Almeida, 2003) and by hiring scientists (learning-by-hiring), or by sending firms scientist abroad to the subsidiaries (Kim et al., 2009). Evidence from the international business literature suggests that knowledge outflows from the multinational corporation's home base are outweighed by inflows from its foreign-based subsidiaries (Singh, 2007; Kogut & Zander, 1993; Dunning, 1992), and that both knowledge flows appear to track personnel flows (Singh, 2007). Focusing on the location of inventor is not a novel way to map the geographical distribution of a firm's innovation activities (Cantwell, 1989) but has attracted less attention than it deserves due to data limitations (Harhoff & Thoma, 2010).

Inventor mobility data can be used to measure the geographical distribution of knowledge spillovers (Breschi & Lissoni, 2009; Almeida & Kogut, 1999; Kim et al., 2009; Agrawal et al., 2006). Mobile individuals are endowed carriers of knowledge stock and play a key role in the diffusion of knowledge by acting as vehicles for knowledge spilloves across organizations and locations through person-to-person interaction. The role of individuals as active agents in the creation and spatial diffusion of knowledge is often emphasized in the literature (Almeida & Kogut, 1999; Howells, 2012), particularly because person-to-person contact involving a transfer or exchange of personnel is gathered as an efficient

---

[5]Beyond mobile workers in the strict sense, *i.e.* workers switching employer or the establishment they work in, mobile inventors can be also consultants or academic scientists that offer their services to different companies (Breschi & Lissoni, 2009).

means of transmission across organizational boundaries for tacit knowledge (Kim et al., 2009). For example, Breschi & Lissoni (2009) argue that the most fundamental reason why geography matters in constraining the diffusion of knowledge is that mobile researchers are not likely to relocate in space, that account to a large extent for localization of co-inventions and citations.

3 Data and Methodology

3.1 Data

The data analysed in this study are drawn from the OECD REGPAT database (Maraut et al., 2008; Webb et al., 2005) which compiles all patent applications filed with the European Patent Office (EPO) from the 1960s to present. In this database the geographical location of each inventor and applicant has been matched to the appropriate 5,552 NUTS3 region in one of the 50 OECD or OECD-partner countries. This allows us to construct 4 geographical networks: (1) co-inventors, (2) applicant-inventor, (3) citations and (4) inventor mobility. For each network we define $y_{m,n}$ as the number of links between NUTS3 region m and n. In (1) $y_{m,n}$ is equal to the number of patents jointly invented by the two regions. We use a full-counting approach so that a patent with $N(>1)$ inventors accounts for $\sum_{i=1}^{N-1}(N-i)$ regional links (hence, patents with only one inventor do not appear in this network by construction). Unlike (1), networks (2), (3) and (4) are directed networks in which we distinguish the pair (m,n) with respect to the pair (n,m). In (2) the region of the applicant is linked to the regions of the affiliated inventors. The inventor's region usually indicates where the invention was made (often a laboratory or a research establishment, or the place of residence of the inventor) while the applicant's region indicates where the holder (usually a company, university or other type of entity) has its headquarters. In the database there is no direct information on affiliations, but it can be trivially retrieved for patents associated with a single applicant, the case for approximately 94% of the whole set of patents. In (3) for each pair (m,n) of NUTS3 regions we count the number of times that (a patent of an inventor in) region m cites (a patent of an inventor in) region n ($y_{m,n}$), and the number of citations that m receives from n ($y_{n,m}$). In (4) a link indicates one inventor moving from one region to another one. Inventors regional migration can be tracked observing patent activity in at least two different years. In the case that an inventor has no patents for one or more years, we can track her region only at the beginning and at the end of the gap. In that case the flow is referred to the first year in which the inventor is observed again.[6] Names of inventors have been cleaned and ambiguity over first names and initials have been dealt with, but have not been fully disambiguated.[7] This results in observing a 13% of inventors who have been active on more than one NUTS3 region in the period 1981-2010.[8]

For the econometric analysis we create a balanced panel of data by networks for the period 1986-2009. The sample used in the estimation is restricted to those pairs for which at least one link is registered in the time period.

---

[6]We stress that, while there might be some overlapping in the mobility and applicant-inventor networks, these capture very different R&D relations between regions. The applicant-inventor network captures the way in which applicant institutions organize the geographical structure of their laboratories. Any inventor move is associated to two applicant-inventor links, the one referring to the outgoing region and the other to the destination region. A data point for inventor moves can correspond to a data point in the applicant-inventor network only for the move destination region and in the particular case that the outgoing region is the same of the region of new applicant. This happens when an applicant relocate the inventor far from the applicant region.

[7]More precisely, our approach tracks the flow of names between regions. As worst, this is a proxy of the flow of individuals, at best, it can get rather close to the real flows. There will be some ambiguity about the source and destination of the move only in the unlikely situation in which two authors with the same name move simultaneously. Our goal is to count the number of moves between regions, not to track the careers of individual inventors over time.

[8]This number is in line with the 9% that Breschi & Lissoni (2009) find analyzing moves of US inventors among Metropolitan Statistical Areas. They focus on the period 1978-2002 though and on a subset of technological fields (Organic Chemistry, Pharmaceuticals and Biotechnology).

3.2 Methodology

The root of our econometric approach is the gravity model (Anderson & Van Wincoop, 2003; McCallum, 1995), a standard tool in the econometrics of trade which has been recently applied to the analysis of R&D networks (Ponds et al., 2007; Scherngell & Barber, 2009; Frenken et al., 2009b; Hoekman et al., 2009, 2010; Scherngell & Barber, 2011; Scherngell & Hu, 2011; Hoekman et al., 2013; Scherngell & Lata, 2012)).[9] The count of links ($y_i \equiv y_{(m,n)}$) between NUTS3 regions (m and n) is regressed on a set of controls which account as a minimum for the geographical distance and for the size of regions. A further set of variables controlling for separation effects is typically added.

We model the dependent variable with a count density. A number of models can be found in the literature to handle count densities, including the Poisson model, Negative Binomial model variants, and Zero-inflated models (Ponds et al., 2007; Scherngell & Barber, 2009; Frenken et al., 2009b; Hoekman et al., 2009, 2010; Scherngell & Barber, 2011; Scherngell & Hu, 2011; Hoekman et al., 2013; Scherngell & Lata, 2012). Since a large portion of NUTS3 region pairs have zero links, we opted for a Zero-Inflated Negative Binomial (ZINB) density, as consistent with Hoekman et al. (2009) and Frenken et al. (2009b). Zero-inflated models allow zeros to be generated by two distinct processes and are generally used when data exhibits "excess zeros" (Cameron & Trivedi, 1998).[10] The ZINB model supplements a count density, $\tilde{P}$, with a binary zero generating process ψ. This allows a zero count to be produced in two ways, either as an outcome of the zero generating process with probability ψ_i, or as an outcome of the count process $\tilde{P}$ provided the zero generating process did not produce a zero ($\psi_i = 1$).

The density distribution for the pair count y_i is then given by

$$P(y_i) = (1 - \psi_i) * \tilde{P}(y_i), \tag{1}$$

where the zero generating process ψ_i is parameterized as a logistic function of the regressors in Z_i, with parameter vector β^0:

$$\psi_i = \frac{\exp(Z_i\beta^0)}{1 + \exp(Z_i\beta^0)}. \tag{2}$$

The count process $\tilde{P}(y_i)$ is modeled as Negative Binomial of the second kind (NB2):

$$\tilde{P}(y_i) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i + 1) + \Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i}\right)^{\alpha^{-1}} \left(\frac{\mu_i}{\alpha^{-1} + \mu_i}\right)^{y_i}, \tag{3}$$

where the conditional mean μ_i is parameterized as an exponential function of the linear index $X\beta^1$ ($\mu_i = \exp(X_i\beta^1)$), and $\alpha \geq 0$ is the overdispersion parameter. Thus, drawing together equations 1, 2, and 3, and assuming $X_i = Z_i$[11] our model for the expected count is

$$E(y_i|X_i) = \left(1 - \frac{\exp(X_i\beta^0)}{1 + \exp(X_i\beta^0)}\right) * \exp(X_i\beta^1) =$$
$$= \frac{\exp(X_i\beta^1)}{1 + \exp(X_i\beta^0)}. \tag{4}$$

The linear indices $X\beta^0$ for the zero-generating process and $X\beta^1$ for the Negative Binomial process

---

[9]See Anderson & Van Wincoop (2004) and Bergrstrand & Egger (2011) for two excellent surveys on gravity model applications. See Anderson (2011) for an updated review on theory.

[10]"Excess zeros" refers to observing more zero observations than what expected with the Poisson distribution.

[11]In our estimation procedure we assume $X_i = Z_i$ because there is no reason to expect some variables would be relevant only in one of the two processes. However, individual regressors can impact the y_i estimator differently through the two distinct processes and their separate parameter vectors, β^0 and β^1.

are modeled in parallel as

$$X\beta^j = \beta_0^j + \beta_1^j border + \beta_2^j distance + \beta_3^j techdist_t + \beta_4^j distance * techdist_t + \beta_5^j neighbour$$
$$+ \beta_6^j size_{m,t} + \beta_7^j size_{n,t} + \sum_{s=2}^{S} \gamma_s^j area_s + \sum_{t=2}^{T} \theta_t^j year_t + \sum_{t=2}^{T} \delta_t^j distance * year_t, \tag{5}$$

where $j = 0, 1$. In this baseline specification we account for five different spatial measures: *border*, *distance*, *techdist*, *neighbour*, *area*. The dummy variable *border* flags pairs of NUTS3 belonging to different countries. The continuous variable *distance* measures the distance, in kilometers, between the centroids of the NUTS3 regions. The continuous variable *techdist* is a measure of the technological distance in a given year.[12] This measure is constructed using patent classes according to the International Patent Classification (IPC). In particular, for each region m we compute the vector $t(m)$ that measures the share of patenting in each of the technological subclasses for a given year. Technological subclasses correspond to the third-digit level of the IPC systems. We define the technological distance between regions m and n as $techdist_{m,n} = 1 - r^2$ where $r^2 = corr[t(m), t(n)]^2$ is the Pearson correlation coefficient between the technological vectors $t(m)$ and $t(n)$ (see Moreno et al., 2005 and Scherngell & Barber, 2009). The dummy variable *neighbour* flag pairs of adjacent NUTS3. The categorical variable *area* splits the network in three kinds of links ($S = 3$) according to the geographical area: links within the EU area, links within the non-EU area and the flows between the two areas. $Size_m$ and $Size_n$ denote the size of each of the two regions. We proxy the size of a region by the total number of links attached to the region. $year_t$ is the year dummy variable.

We make use of the general model highlighted in equations 4 and 5 to perform three sets of estimates according to our research questions. For the first and second cases we run estimates on a balanced panel of data by networks for the period 1988-2009. The sample used for estimation is made of all pairs of regions with at least 20 patents in every year. We chose to set a threshold on patents for two reasons. First, the large majority of NUTS3 regions pairs have no links, which are concentrated on inter-regional pairs with few patents. Second, our measure of technological distance requires a reasonable amount of patents to be reliable.[13]

To test our first research question, we make use of maximum likelihood estimates of parameters in equation 5 and compute the elasticity of *distance* over years. Specifically we estimate for each year the quantity $\epsilon_t = \frac{\partial E(y_i)}{\partial distance} \frac{distance}{E(y_i)}$.[14]

To test our second question we estimate the evolution of the country-border effect. To do this we modify equation 5, adding interactions of *border* with year dummies, resulting in

$$X\beta^j = \beta_0^j + \beta_1^j border + \beta_2^j distance + \beta_3^j techdist_t + \beta_4^j distance * techdist_t$$
$$+ \beta_6^j size_{m,t} + \beta_7^j size_{n,t} + \sum_{s=2}^{S} \gamma_s^j area_s + \sum_{t=2}^{T} \theta_t^j year_t + \sum_{t=2}^{T} \delta_t^j distance * year_t$$
$$+ \sum_{t=2}^{T} \omega_t^j border * year_t. \tag{6}$$

Given maximum likelihood estimates of parameters in the augmented equation 6[15] we compute the

---

[12]The interaction term $distance\_techdist_t$ is also included. Estimates of elasticity for distance over different levels of *techdist* can be provided upon request.

[13]Robustness checks were performed using different thresholds, both lower and higher than 20. Results hold very similar to those reported in this article. These are made available by the authors.

[14]From equation 4 we compute the derivative as $\frac{\partial E(y_i)}{\partial distance} = E(y_i) \left(\beta_2^1 - \beta_2^0 \frac{\exp(X_i\beta^0)}{1+\exp(X_i\beta^0)} \right)$. Thus we have $\epsilon_t = distance \left(\beta_2^1 - \beta_2^0 \frac{\exp(X_i\beta^0)}{1+\exp(X_i\beta^0)} \right)$. Estimates of this quantity are obtained replacing parameter estimates and setting sample mean values for regressors.

[15]In equation 6 we omit the dummy *neighbour* as it gets highly collinear and renders maximum likelihood convergence

marginal effects of the *border* variable over years. In particular we compute for each year the quantity $\Delta_t = E\left(y_i|, border = 1, year_t = 1, X_i = \bar{X}\right) - E\left(y_i|, border = 0, year_t = 1, X_i = \bar{X}\right)$.[16] Then we relativize this difference to $E\left(y_i|, border = 0, year_t = 1, X_i = \bar{X}\right)$ obtaining the semi-elasticity, the quantity we report in the results section to allow comparisons among networks.

The estimates for investigating our third question are performed by changing the specification of equation 5 and separating the data into EU and non-EU sets. The sample used in the estimation is a balanced panel of data for the period 1986-2009. The sample is restricted to those pairs for which at least one link is registered in the time period.[17] As regards the specification of the linear indices, we stick to Chessa et al. (2013) in applying a Difference-in-Difference (DiD) strategy to isolate the country border effect within EU.[18] The rate at which EU (NUTS3) regions are linking to regions in other EU countries is increasing due to two types of factors: those that are global and those that are EU specific. Thus, to capture the effect of EU specific institutional factors we must account for the net effect of the global factors. In technical terms, we use the non-EU OECD members as a control group and its behaviour serves as the counterfactual behavior of EU regions.[19] The linear indices in 5 are now modeled as

$$X\beta^j = \beta_0^j + \beta_1^j border + \beta_2^j eu + \beta_3^j distance + \beta_4^j size_m + \beta_5^j size_n + \gamma^j border * eu +$$
$$+ \sum_{t=2}^T \theta_t^j year_t + \sum_{t=2}^T \delta_t^j border * year_t + \sum_{t=2}^T \zeta_t^j eu * year_t + \sum_{t=2}^T \eta_t^j border * eu * year_t, \qquad (7)$$

where the trinomial variable *area* collapses in the binomial variable *eu* as links between the EU area and non-EU area are removed for identification purpose. In particular, *eu* flags pairs of NUTS3 regions that are within the EU ($eu = 1$) and pairs of NUTS3 regions for which neither are in the EU ($eu = 0$). *border* still flags pairs of NUTS3 regions within the same country but now links pertain always to the same area (EU or non-EU) whether or not they are cross-border or within-border.[20] In terms of the standard DiD formalism (Angrist & Krueger, 1999; Heckman et al., 1999; Athey & Imbens, 2006; Blundell & Costa Dias, 2009) the dummy border distinguish treated units, *i.e.* regions belonging to different countries, from un-treated units, *i.e.* regions within the same country. Then to isolate the signal arising only from EU factors we extend the standard DiD strategy of one state indicator (treatment vs control group) to the case of two state indicators, providing a further control group of links between non-European countries. For the purpose of embedding the institutional comparison in a temporal perspective, our analysis also includes year dummies. Due to the addition of a second state indicator our approach is a Difference-in-Differences-in-Differences estimator (DiDiD). The full set of double/triple interaction dummy variables among the three dimensions ($eu = \{0,1\}$, $border = \{0,1\}$, $year_t = \{0,1\}$ for $t = 2, \ldots, T$) is relevant to the identification of treatment effect (Wooldridge, 2010). In this framework, treatment effects are incremental effects of the triple interaction terms $border * eu * year_t$.

Denoting the actual and counterfactual outcomes of our count dependent variable as y_i^T and y_i^C respectively and taking into account our DiDiD extension, the yearly treatment effect (τ_t) can be

---

cumbersome.

[16]The quantities of the difference can be easily retrieved from equation 4 and 6 using parameter estimates and replacing $border = 1$ or $border = 0$, $year_t = 1$ and sample means for regressors. See Winkelmann (2008) for the computation of marginal effects for the ZINB model.

[17]Here we do not include in the regressions the variable *techdist* which require to set a threshold on the number of patents.

[18]See Chessa et al. (2013) for a more detailed description of the methodology.

[19]EU recent members are removed from the group of non-EU OECD members.

[20]For example, Italy-France can be a valid cross-border link for EU and USA-Japan can be a valid cross-border link for non-EU. However Italy-USA, Italy-Japan, France-USA and France-Japan are excluded. Such links are simply not relevant to the comparison we are focusing on.

defined as

$$
\begin{aligned}
\tau_t(year_t = 1, border = 1, eu = 1, M) = &E\left(y^T | year_t = 1, border = 1, eu = 1, M\right) \\
&- E\left(y^C | year_t = 1, border = 1, eu = 1, M\right),
\end{aligned}
\tag{8}
$$

where M is the matrix of controls $(Size_m, Size_n, Distance)$. The quantities in equation 8 can be easily computed replacing in equation 4 maximum likelihood estimates of parameters in equation 7 and specific values for regressors. We refer all values to a generic pair of cross-border EU regions in the baseline year. Relative to the baseline year t^* (we use the arbitrarily chosen year 2004), the yearly treatment effect reflects the impact of changes in institutional factors specific to the EU which have taken place in a given year t with respect to t^*. Nevertheless, because differences in the treatment effect relative to the baseline year are transitive, trends across time, and in particular between years that do not include the baseline year, can still be interpreted as gross differences.

4   Results

In this section we present the results of our three regression estimates about the evolution of the distance and border effect, and the relative evolution of the cross-border effect in Europe.

4.1   Evolution of the distance effect

Figure 1 shows the evolution of the average distance of R&D collaborations between NUTS3 regions. Clearly, distance is more important when the flow of knowledge is based on human interactions (*i.e.* co-inventor, applicant-inventor) or mobility. While these networks require costs of moving or costs of communication, citations on the other hand benefit from the availability of online repositories of bibliometric records. Moreover, as expected and documented in the literature, the average distance of R&D collaboration grows over time. This is true for each network, with mobility standing out for its strongest trend. The average distance of inventor moves was 1,267 kilometers in 1986 and has almost doubled by 2009 reaching 2,051 kilometers. However, one should consider that this summary statistics do not take into account inter-regional links.

In Figure 2 we show the results of our econometric analysis, specifically network-by-network estimates of the elasticity of $y_{m,n}$ with respect to distance. For each year we report the point elasticity evaluated at sample means and the 95% confidence interval. For example in 2008, taken an "average" pair of NUTS3 regions, a 1% increase in the distance implies a decrease in the number of links of 1.24% for applicant-inventor, 1.10% for co-inventor, 0.94% for mobility and 0.27% for citations. These estimates show distance clearly is still a major constraint of inter-regional connectivity for every network, and the citation network result confirm distance impedes the flow of codified knowledge much less than tacit.

Looking at the time evolution of the distance effect we observe a positive trend for co-inventor, applicant-inventor and citations. For the first two, the positive trend emerges in the early Nineties, while for citations it starts earlier and is roughly stable from 2002 onwards (see Figure 1).[21] For mobility, the effect of distance exhibits a significant decrease in earlier years reaching the minimum in 1997, but then stabilizes.

Evidence that the effect of distance is increasing over time in three of the networks is apparently at odds with the earlier observation that average distance of inter-regional links increases over time. This is a general pattern of developed countries and is not being driven by specific countries or group

---

[21]The temporal evolution of the distance effect holds very similar making use of the augmented equation 6, *i.e.* when also the temporal evolution of the cross-border effect is accounted for. Results are available upon request.

of similar countries. As can be noticed in Figure 3, for the co-inventor network the same results are found for a large core European country (Germany), the European Union (EU15), right down to a group of small countries in the core of the network (Belgium, Netherlands, Luxembourg, Austria and Switzerland) and a group of inter-connected countries in the European periphery (Norway, Sweden, Finland, Denmark and Island). Further, a positive trend persists in estimates even when we remove regressors one by one until we end up with the basic gravity model that includes only physical distance and the size of nodes. It is crucial to note that the trend in the effect of distance turns negative only once size of nodes (regions) is omitted from the regression, leaving us with a dependent variable that is regressed only against distance and interaction with the year dummy. Figure 4 shows this reversal in the trend for the applicant-inventor case. The insight we can draw from this Figure is that the increase in the average distance of collaboration can be explained by the leading regions reaching-out, *i.e.* the attractiveness of regions with large number of connections. As the central nodes grow and grow, peripheral nodes are more likely to connect to them, resulting in an increase of the average distance. However, once we clean the pulling force between two nodes of the role played by their sizes using a gravity approach, what is left is a time-increasing distance effect. We note also that by reintroducing the other controls the trend in the effect of distance remaining similar.

4.2 Evolution of the border effect

In Figure 5 we report the evolution of the border effect expressed as semi-elasticity, *i.e.* the percentage change in the count of links when the dummy *border* shifts from 0 to 1. For example in 2008, taking an "average" pair of NUTS3 regions, the country border reduces the number of links by 92.9% for co-inventor, 92.5% for applicant-inventor, 87.3% for mobility and 59.2% for citations. The effect of country borders is clearly quite strong. Similar to the distance effect, it is far less important for the citation network and more important in co-inventor and applicant-inventor than in mobility.

Unlike the distance effect, we find some sign that the border effect is decreasing for the co-inventor, applicant-inventor and citation networks. For those three networks the trend is overall negative, though only mildly and with periods of positive trend. Notably we observe an increase in the border effect starting in 1996 for co-inventor and 1997 for applicant-inventor and citations, until recent years, though levels still remain lower than 1988. For mobility, the border effect resembles what we saw for the distance effect, with significant decline until 1995 and a flattening afterwards.

To provide a better understanding of the kind of collaborations which happen more often across countries, we split regional links according to the size of regions. We identify the 100 top regions as those having the highest number of patents filed in the period 1986-2009 and distinguish three groups of links respectively between (1) large and large regions, (2) large and small regions, (3) small and small regions.[22] In Figure 9 we report the temporal evolution of the cross-border share of co-inventor links for these three groups. An increase in the share of cross-border links is a common feature. Looking at the levels we notice that the cross-border share is typically larger for links between small and large regions suggesting that when small regions collaborate across borders they are relatively more likely to collaborate with large regions. This share is also the fastest increasing, suggesting that, in the light of the regressions' results, an easing of the cross-border effect can be at least in part explained by the higher ability of small regions to match large regions. Links between top regions happen rarely across borders instead, with only a mild increase over the period we examine.[23]

---

[22]The percentages of links pertaining to the three groups are 44.1% (Top100-Top100), 18.2% (Top100-NonTop100), 37.7% (NonTop100-NonTop100) respectively.

[23]A similar reaching out effect has been already noticed for the US Life Sciences patent network by Owen-Smith et al. (2002)

4.3 Evolution of the European research integration

Lumping together all OECD countries, the results presented in the previous section on the evolution of the border effect tells us nothing about the effectiveness of the largest cross-border integration effort, the European policies undertaken to stimulate integration in R&D in the EU. Hence, in order to measure the role of borders in EU vis-à-vis non-EU collaboration networks, we perform a comparative analysis. Figure 7 shows the yearly treatment effects which quantify the relative impact of EU-specific factors on cross-border connectivity. Indeed, since the late Nineties, we observe some positive signs of integration in patent statistics. In the case of the co-inventor network, we find an increasing trend of cross-border collaboration between inventors in Europe vis-à-vis other OECD countries. This effect was relatively pronounced after 2000, the launch year of ERA initiative, but has stalled since 2005. This partial integration of the network of inventors has not been complemented by an analogous trend in the other networks. Apart some positive effects in the late Eighties/early Nineties, no significant trends can be depicted from early Nineties to the end of the sample period.[24]

The analysis of the Herfindahl index of technological specialization of regions shows that there is a global trend of increasing technological concentration over time. This tendency affects large and small regions as well, with large regions which stay on average more diversified than the small ones. The growing specialization of regions reflects into the increase of the average technological distance among regions. Against this background, Figure 8 shows that top regions outside Europe tend to become relatively more specialized than European counterparts. This trend suggest that the European strategy for smart specialization should focus on augmenting the technological focus of European clusters.

To look deeper at the geographic heterogeneity of European integration we calculate the rate at which each EU region is integrating with other EU regions, and the rate at which each is integrating with non-European regions. Specifically, over the period 1986 to 2009, we calculate for each EU region its slope on: (a) percentage of links that cross borders to other EU regions; (b) percentage of links that cross borders to regions outside of Europe. Figure 6 is a graphic representation of these results for the co-inventor network. Panel (a) depicts the rate of integration within the EU and (b) the rate of integration to regions outside of Europe. The colour scale runs from light yellow to dark red. The gradient scale is the same for both panels, is linear, and starts at zero. It is important to note that the scale minimum is zero as we simply did not observe a single EU region with a negative integration rate (calculated as defined earlier in this paragraph). Noting that, in general, colour intensities are higher in Figure 6(a) than in Figure 6(b) we can conclude that intra-EU integration rates are typically higher than extra-European. However it is important to keep in mind that it is non-trivial to relate this observation to the regression results presented earlier because there we focused on integration *within* the EU and integration *among* non-European countries, there by ignoring integration between Europe and the rest of the world (*i.e.* the data in Figure 6(b)). In Figure 6(a) it appears that Middle European regions close to the German border, and have been highlighted as insets. In the case of Benelux the regions experiencing the largest increase in integration seem to lie on the borders. As noted above, for Figure 6(b) values are typically lower, however the UK and Ireland also stand out "hotspots" of integration with the rest of the world. Indeed, within the UK and Ireland there are several regions who's average rate of extra-Europe integration is even greater than their intra-EU rate.

---

[24]In the case of inventor mobility the number of non-zero link counts was too low to be modeled using ZINB, thus estimation was carried out aggregating the network at NUTS2 level. We replicated the analysis including also New Member States in the group of European NUTS3 regions. In the augmented group of 27 countries the role of new members in the R&D networks is anyways very small, accounting for a tiny percentage of the whole links. This is reflected in estimates as the evolution of the treatment effects is very similar to what we reported.

5   Discussion

We analzyed the temporal evolution of spatial biases in the strength of inter-regional connectivity within a set of regional patent networks. Focusing on a set of fifty developed nations over the period 1988-2009, using inter-regional links at the NUTS3 region level, we have contributed to the body of literature on the geography of knowledge by analyzing jointly a set of four R&D networks: co-inventor, applicant-inventor, citations, inventor mobility. Making use of a gravity-like econometric approach and controlling for a number of separation effects, we estimated year-by-year effects of physical distance and country-borders, and the trend of integration in the European Research Area.

Contrary to the widespread notion that the importance of distance has been decreasing over time due to globalization and technological advancement, estimates reported in Section 4 show that the constraint imposed by geographical distance on R&D inter-regional links seem to have actually increased in three of the networks analyzed: co-inventor, applicant-inventor, and citations. On average, inter-regional links take place at a larger distance, which can be intuitively understood as large nodes increasing their attractiveness to peripheral nodes as they grow in time. However, *ceteris paribus*, for a pair of regions of a given size the strength of their connectivity gets more sensitive to physical distance with time. This means the cost of inter-regional collaboration at a given distance is still large, even increasing, but whenever a small region becomes connected to a hub the relevance of this cost is counterbalanced by the benefit of linking to a core region. Indeed, large and diversified regions tend to extend their basin of attraction across national borders, prevalently toward small regions. This trend is a driving force of our estimates for the evolution of the cross-border effect, which indicates that national borders can be crossed more easily now than in the late 1980s, particularly due to a significant decrease up to the mid 1990s. Hence, when we observe collaborations to happen more frequently across borders, this is largely driven both by an erosion on institutional frictions that impede inter-national connectivity (Hoekman et al., 2010) and the "reaching-out" by international hubs, rather than a decrease in the costs associated with collaborating over distance. Part of this story can be explained in terms of the role played by the European Union in promoting inter-national connectivity within the area, though signs of integration are weak an limited to collaboration between inventors.

In estimating the evolution of the distance effect we note that the mobility network stands out as the only network with a negative trend, though with no significant change after 1997. This suggests that the globalization of skilled-labor job markets which enabled a reduction in mobility costs have had a larger impact on the geography of knowledge than advances that favour a reduction in cost of communication. In particular, the result that the distance effect is steadily increasing in the network of citations despite well-known advances in technologies easing the codification of knowledge corroborates the notion that tacit and embodied knowledge still play a major role in diffusion. In particular, patents are pieces of codified knowledge building upon a stock of tacit knowledge that hinders its fruition (Breschi & Lissoni, 2009). Overall, the increase in distance effect supports the view that improvements in communication technologies, while on the one hand facilitating the substitution of face-to-face interactions with arm's-length communication, on the other hand create a greater need for close interactions to exchange complex knowledge which is responsible for research activities to agglomerate rather than to disperse (Gaspar & Glaeser, 1998).

For the geographical dispersion of the network of R&D activities we observe a significant decrease of the border effect over the period 1988-1997 and only a mild decline of the distance effect over a similar time window (1988-1996). We also notice that the distance effect and the border effect increase almost hand-in-hand since the late 1990s. An increase in these effects means that, once we account for the effect of size and other variables, inventors are more likely to be located nearby and in the

same country as their institution (patent applicant). Excessive geographical dispersion of learning centers can lead to difficulties in controlling the generation and exploitation of knowledge, especially given its predominant content of tacitness. This argument has been invoked to explain a substantial change in international location decisions observed immediately after an opposite trend between 1985 and 1995. In fact, the strong movement to establish a transnational configuration of R&D observed between 1985 and 1995 has been blamed to result in overly complex and unmanageable organizational architectures (Gerybadze & Reger, 1999). In the light of the role played by knowledge-seeking reasons in the internalization of innovative activities, our results are coherent with these trends and point out that limits encountered by R&D internalization strategies in controlling the accumulation of knowledge across geographical and institutional borders have not been reduced by globalization forces.

We note that the time window where we observe a significant decrease in the border effect for mobility is somehow related to the period of decreasing border effect in the other R&D networks. This applies also for the distance effect, at least for co-inventorship and the location of R&D activities. Thus our results reinforce the view that individual mobility is the driving force of knowledge integration (Breschi & Lissoni, 2009).

As concluding remarks concerning policy, we stress the importance of R&D clusters. Our evidence suggest that integration in research is being driven by the top regions reaching out to more peripheral regions and across borders. This trend in the evolution of R&D networks supports policies oriented to the exploitation of agglomeration economies in research clusters rather than targeting promotion of cross-border collaboration (Hoekman et al., 2010). This trend is in line with smart specialization strategies as they can be a valuable asset to speed up the creation and consolidation of a European Research Area. However, the importance of investment in programs that incentivize mobility of researchers throughout Europe seems to be reaffirmed, even if we do not have explicit evidence of tangible benefits in the European Union as opposed to the rest of the developed world.

As a final remark, we point out some limitations in our analysis, which could be addressed in future research. We do not consider scientific publications or R&D projects and collaborative agreements in our analysis. Further investigation is needed to assess whether similar trends are present for basic research and other networks of innovators. Another extension maybe to explicitly test for the dynamic interplay between different R&D networks. Finally, the increasing availability of large data sets of bibliometric information should encourage the application of new quantitative methods to assess the efficacy of the European R&D policies for smart specialization and integration.

References

A. Agrawal, et al. (2006). 'Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships'. *Journal of Economic Geography* **6**:571–591.

A. Agrawal & A. Goldfarb (2008). 'Restructuring Research: Communication Costs and the Democratization of University Innovation'. *American Economic Review* **98**(4):1578–1590.

P. Almeida & B. Kogut (1999). 'Localisation of knowledge and the mobility of engineers in regional networks'. *Management Science* **45**(7):905–917.

J. E. Anderson (2011). 'The gravity model'. *Annual Reviews of Economics* **3**:133–160.

J. E. Anderson & E. Van Wincoop (2003). 'Gravity with gravitas: A solution to the border puzzle'. *American Economic Review* **93**(1):170–192.

J. E. Anderson & E. Van Wincoop (2004). 'Trade costs'. *Journal of Economic Literature* **42**(3):691–751.

J. Angrist & A. Krueger (1999). 'Empirical Strategies in Labor Economics'. In O. Ashenfelter & D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, pp. 1277–1366. Elsevier.

S. Athey & G. Imbens (2006). 'Identification and inference in nonlinear difference-in-differences models'. *Econometrica* **74**(2):431–497.

G. H. Bergrstrand & P. Egger (2011). 'Gravity equations and economic frictions in the world economy'. In D. Bernhofen, R. Falvey, D. Greenaway, & U. Krieckemeier (eds.), *Palgrave Handbook of International Trade*. Palgrave Macmillan Press, New York.

R. Blundell & M. Costa Dias (2009). 'Alternative Approaches to Evaluation in Empirical Microeconomics'. *Journal of Human Resources* **44**(3):565–640.

Boerner, et al. (2006). 'Mapping the diffusion of scholarly knowledge among major US research institutions'. *Scientometrics* **68**(3):415–426.

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

S. Breschi & C. Malerba (2005). *Clusters, Networks, and innovation*. Oxford University Press, Oxford.

J. Brun, et al. (2005). 'Has distance died? Evidence from a panel gravity model'. *World Bank Economic Review* **19**(1):99–120.

F. Cairncross (1997). *The Death of Distance: How the Communications Revolution will Change Our Lives*. Harvard Business School Publishing, Cambridge, MA.

A. C. Cameron & P. K. Trivedi (1998). *Regression analysis of count data*. Cambridge University Press, Cambridge.

J. Cantwell (1989). *Technological Innovation and Multinational Corporations*. Basil Blackwell, Oxford.

M. Castells (1996). *The Rise of the Network Society*. Blackwell Publishers, Oxford.

A. Chessa, et al. (2013). 'Is Europe evolving toward an integrated research area?'. *Science* **39**(6120):650–651.

D. Coe (2002). 'The missing globalization puzzle'. *IMF Working Paper* (171).

J. H. Dunning (1992). *Multinational enterprises and the global economy*. Addison-Wesley, Wokingham, England.

J. H. Dunning & S. M. Lundan (2009). 'The Internationalisation of Corporate R&D: A Review of the Evidence and Some Policy Implications for Home Countries'. *Review of Policy Research* **26**:13–33.

M. P. Feldman (1994). *The geography of innovation*, vol. 2. Kluwer Academic Publisher.

R. Florida (1997). 'The globalization of R&D: Results of a survey of foreign-affiliated R&D laboratories in the USA'. *Research Policy* **26**:85–103.

D. Foray & B. Van Ark (2007). 'Smart specialisation in a truly integrated research area is the key to attracting more R&D to Europe'. Knowledge Economists Policy Brief 1, Knowledge for Growth Expert Group, European Commission.

C. Freeman (1991). 'Networks of innovators: a synthesis of research issues'. *Research policy* **20**(5):499–514.

C. Freeman (1995). 'The "national system of innovation" in historical perspective'. *Cambridge Journal of Economics* **19**:5–24.

K. Frenken (2002). 'A new indicator of European integration and an application to collaboration in scientific research'. *Economic Systems Research* **14**(4):345–361.

K. Frenken, et al. (2009a). 'Spatial scientometrics: towards a cumulative research program'. *Journal of Informetrics* **3**(3):222–232.

K. Frenken, et al. (2009b). 'Death of Distance in Science? A Gravity Approach to Research Collaboration'. In A. Pyka & A. Scharnhorst (eds.), *innovation Networks*, pp. 43–57. Springer-Verlag, New York.

J. Gaspar & E. L. Glaeser (1998). 'Information Technology and the Future of Cities'. *Journal of Urban economics* **43**:136–156.

A. Gerybadze & G. Reger (1999). 'Globalization of R&D: recent changes in the management of innovation in transnational corporations'. *Research Policy* **28**:251–274.

E. L. Glaeser, et al. (1992). 'Growth in Cities'. *Journal of Political Economy* **100**(6):1126–1152.

B. Gomes-Casseres, et al. (2006). 'Do Alliances Promote Knowledge Flows?'. *Journal of Financial Economics* **80**:5–33.

O. Granstrand (1999). 'Internationalisation of corporate R&D: a study of Japanese and Swedish corporations'. *Research Policy* **28**:275–302.

D. Harhoff & G. Thoma (2010). 'Inventor Location and the Globalization of R&D'.

J. Heckman, et al. (1999). 'The Economics and Econometrics of Active Labor Market Programs'. In O. Ashenfelter & D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, pp. 1865–2097. Elsevier.

J. Hoekman, et al. (2010). 'Research collaboration at a distance: Changing spatial patterns of scientific collaboration in Europe'. *Research Policy* **39**.

J. Hoekman, et al. (2009). 'The geography of collaborative knowledge production in Europe'. *The Annals of Regional Science* **43**:721–738.

J. Hoekman, et al. (2013). 'Acquisition of European research funds and its effect on international scientific collaboration'. *Journal of Economic Geography* **13**(1).

J. Howells (2012). 'The geography of knowledge: never so close but never so far apart'. *Journal of Economic Geography* **12**(5):1003–1020.

A. Jaffe & M. Trajtenberg (2002). *Patents, Citations and Innovations: A Window on the Knowledge Economy*. MIT Press, Cambridge, MA.

A. B. Jaffe, et al. (1993). 'Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations'. *The Quarterly Journal of Economics* **434**:578–598.

W. Keller (2004). 'International Technology Diffusion'. *Journal of Economic Literature* **42**(3):752–782.

J. Kim, et al. (2009). 'International Knowledge Flows: Evidence from an Inventor-Firm Matched Data Set'. In R. B. Freeman & D. L. Goroff (eds.), *Science and Engineering Careers in the United States: An Analysis of Markets and Employment*, chap. 10, pp. 321–348. University of Chicago Press.

B. Kogut & U. Zander (1993). 'Knowledge of the firm and the evolutionary theory of the multinational corporation'. *Journal of International Business Studies* **24**(4):625–645.

S. Long & J. Freese (2006). *Regression Models for Categorical Dependent Variables Using Stata (2nd edition)*. Stata Press, College Station Texas, US.

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

M. A. Maggioni & T. E. Uberti (2007). 'Inter-regional knowledge flows in Europe: An econometric analysis'. In K. Frenken (ed.), *Applied Evolutionary Economics and Economic Geography*, pp. 230–255. Edward Elgar, Cheltenham.

S. Maraut, et al. (2008). 'The OECD REGPAT Database: a Presentation'. Working Paper 2, OECD.

J. McCallum (1995). 'National Borders Matter: Canada–U.S. Regional Trade Patterns'. *American Economic Review* **85**(3):615–623.

R. Moreno, et al. (2005). 'Spatial spillovers and innovation activity in European regions'. *Environment and Planning A* (37):1793–1812.

R. Narula & A. Zanfei (2005). 'Globalization of Innovation: The Role of Multinational Enterprises'. In J. Fagerberg, R. R. Nelson, & D. C. Mowery (eds.), *Handbook of Innovation*, chap. 12, pp. 99–99. Oxford University Press, Oxford.

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

J. Owen-Smith, et al. (2002). 'A Comparison of U.S. and European University-Industry Relations in the Life Sciences'. *Management Science* **48**(24).

M. Ozman (2009). 'Inter-firm networks and innovation: a survey of literature'. *Economic of Innovation and New Technology* **18**(1):39–67.

P. Patel & M. Vega (1999). 'Patterns of Internationalisation of Corporate Technology: Location vs. Home Country Advantages'. *Research Policy* **28**:145–155.

G. Peri (2005). 'Determinants of knowledge flows and their effect on innovation'. *The Review of Economics and Statistics* **87**(2):308–322.

A. Phene & P. Almeida (2003). 'How do firms evolve? The patterns of technological evolution of semiconductor subsidiaries'. *International Business Review* **12**(3):349–367.

R. Ponds (2009). 'The limits to internationalisation of scientific research collaboration'. *Journal of Technology Transfer* **34**:76–94.

R. Ponds, et al. (2007). 'The geographical and institutional proximity of research collaboration'. *Papers in Regional Science* **86**.

W. W. Powell & S. Grodal (2005). 'Networks of Innovators'. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (eds.), *The Oxford handbook of innovation*, chap. 3, pp. 56–85. Oxford University Press, Oxford.

T. Scherngell & M. J. Barber (2009). 'Spatial interaction modelling of cross-region R&D collaborations. Empirical evidence from the 5th EU Framework Programme'. *Papers in Regional Science* **88**:531–645.

T. Scherngell & M. J. Barber (2011). 'Distinct spatial characteristics of industrial and public research collaborations: Evidence from the fifth EU framework programme'. *The Annals of Regional Science* **46**(247).

T. Scherngell & Y. Hu (2011). 'Collaborative knowledge production in China. Regional Evidence from a Gravity Model Approach'. *Regional Studies* **45**(6):755–772.

T. Scherngell & R. Lata (2012). 'Towards an integrated European Research Area? Findings form eigenvector spatially filtered spatial interaction models using European Framework Programme data'. *Papers in Regional Sciences (forthcoming)* .

J. Singh (2007). 'Asymmetry of knowledge spillovers between MNCs and host country firms'. *Journal of International Business Studies* **38**(5):764–786.

J. Singh & M. Marx (2012). 'Geographic Constraints on Knowledge Diffusion: Political Borders vs. Spatial Proximity'. *Working Paper* .

M. von Zedtwitz & O. Gassmann (2002). 'Market versus Technology Drive in R&D Internationalisation: Four Patterns of Managing Research and Development'. *Research Policy* **31**(4):569–588.

Q. H. Vuong (1989). 'Likelihood ratio tests for model selection and non-nested hypotheses'. *Econometrica* **57**.

C. Webb, et al. (2005). 'Analysing European and International Patent Citations: a Set of EPO Patent Database Building Blocks'. Working Paper 9, OECD.

K. B. Whittington, et al. (2009). 'Networks, Propinquity, and Innovation in Knowledge-Intensive Industries'. *Administrative Science Quarterly* **54**(1):90–122.

R. Winkelmann (2008). *Econometric Analysis of Count Data*. Springer-Verlag, Berlin.

Y. Wolfmayr, et al. (2013). 'The role and internationalisation strategies of multinational companies in innovation'. Inno-grips ii report, European Commission, DG Enterprise and Industry, Brussels.

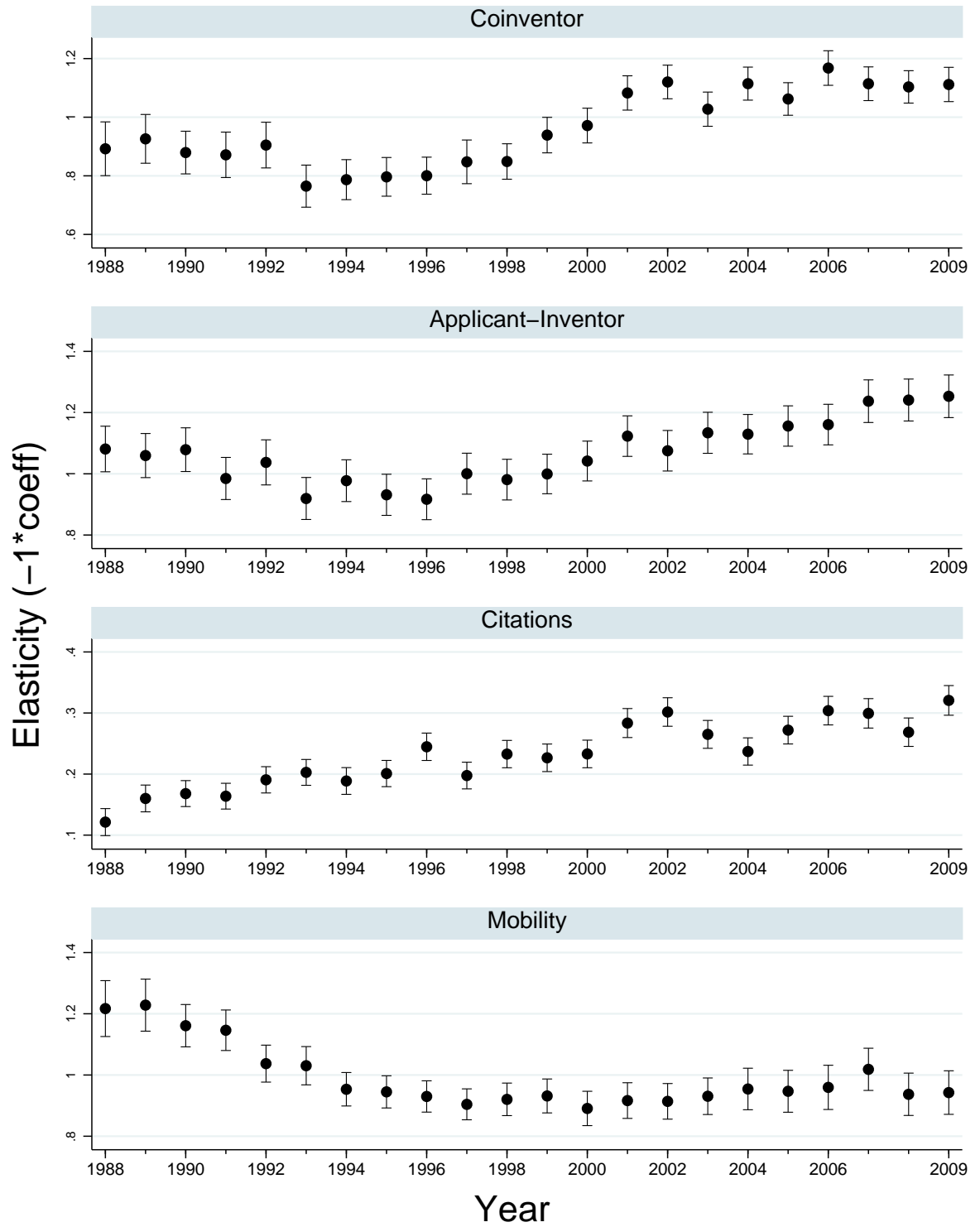J. M. Wooldridge (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge.

Figures
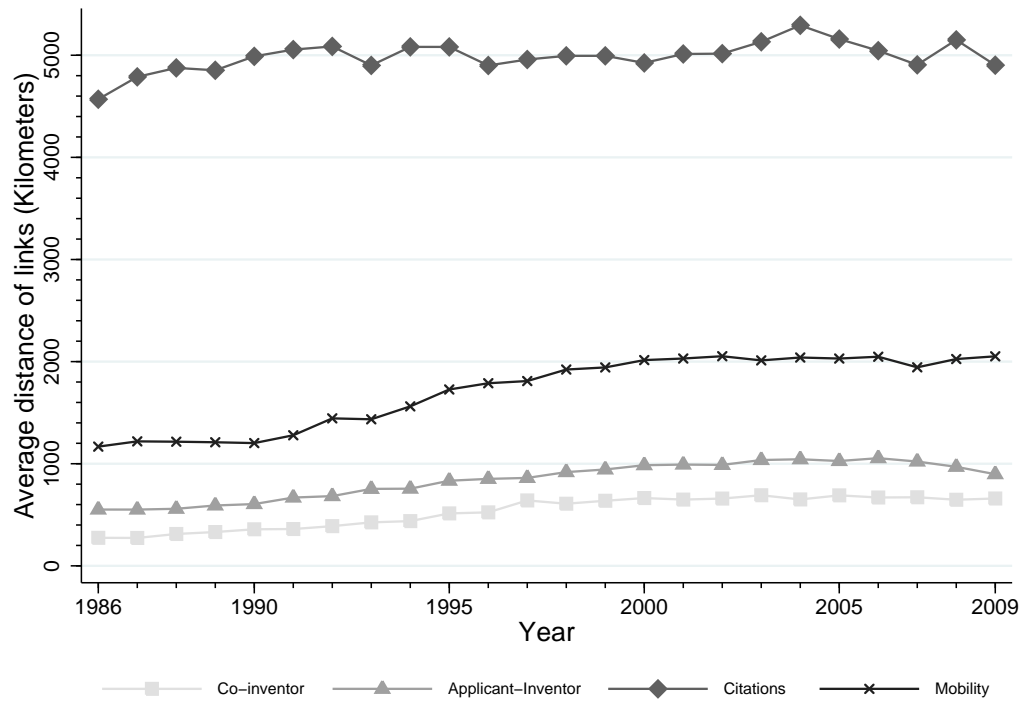
Figure 2
ELASTICITY OF DISTANCE OVER YEARS

Notes: Estimates derive from four separate ZINB-gravity models for the count of links between NUTS3 regions. In the graph we report the year-elasticity and the corresponding 95% confidence interval constructed estimating standard errors through the Delta method. Marginal effects, *i.e.* elasticities, are computed assuming mean values for regressors. We report in Table 1 in the Appendix regression estimates from which elasticities are calculated. Year estimates of the distance effect are obtained including interaction terms between the continuous distance variable and year dummies.

Figure 3

ELASTICITY OF DISTANCE OVER YEARS BY GROUPS OF EU COUNTRIES — CO-INVENTOR

SMALL–CORE: BE, NL, LU, AT, CH; SCANDINAVIA: NO, SE, FI, DK, IS
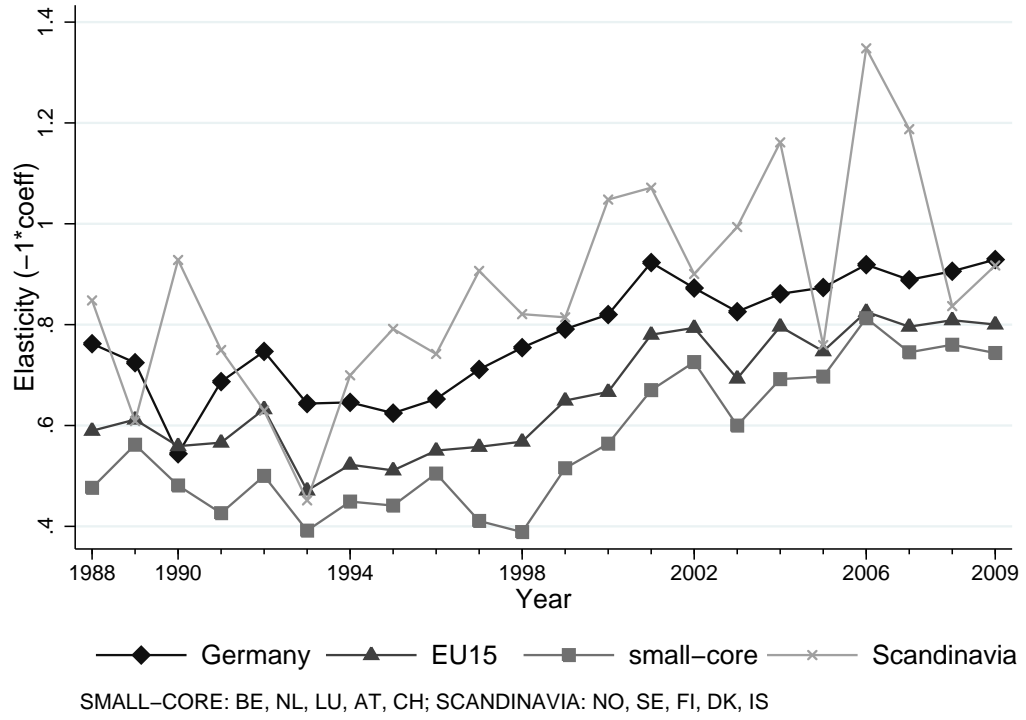
Notes: See notes in Figure 2. For each set of countries we retain in the estimation sample all links pertaining to at least one country in the set.

Figure 4

ELASTICITY OF DISTANCE WITH DIFFERENT MODELS — APPLICANT-INVENTOR

Notes: Each line plot refer to estimates of different ZINB model (see notes to Figure 2).
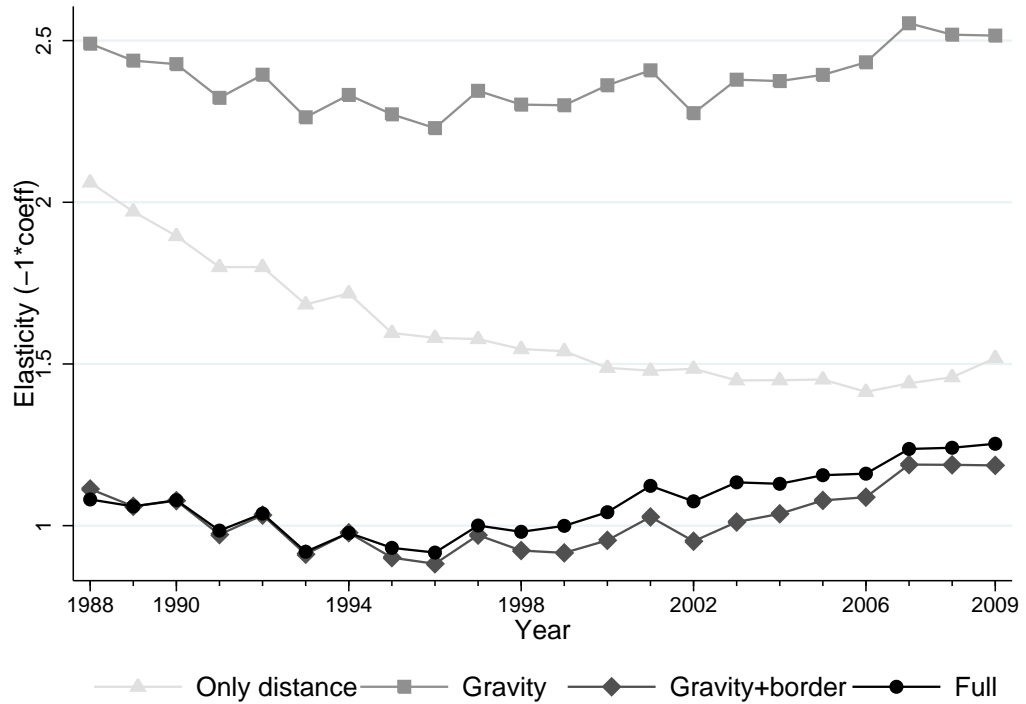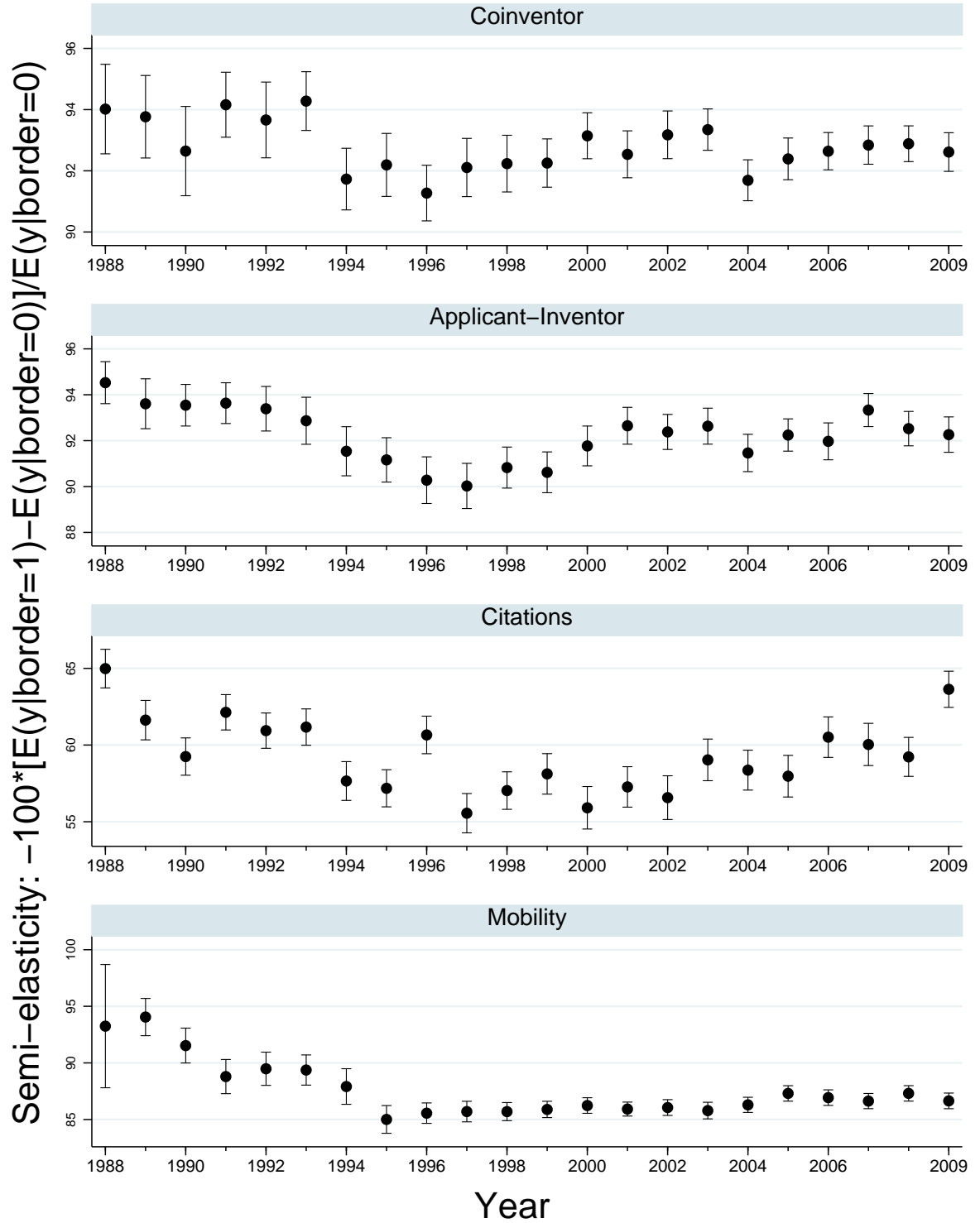
Figure 5
BORDER EFFECT OVER YEARS

Notes: Estimates derive from four separate ZINB-gravity models for the count of links between NUTS3 regions. In the graph we report the yearly semi-elasticity and the corresponding 95% confidence interval constructed estimating standard errors through the Delta method. Marginal effects, *i.e.* semi-elasticities, are computed assuming mean values for regressors. We report in Table 2 in the Appendix regression estimates from which semi-elasticities are calculated. Year estimates of the country-border effect are obtained including interaction terms between the cross-border dummy and year dummies.
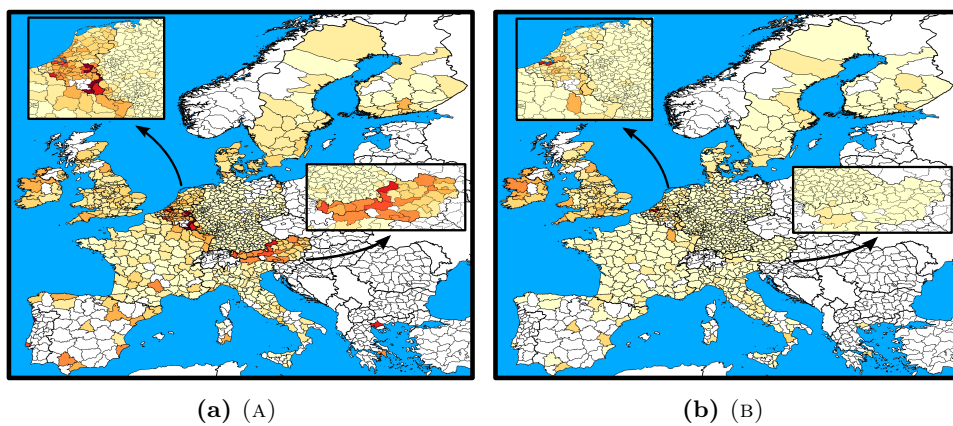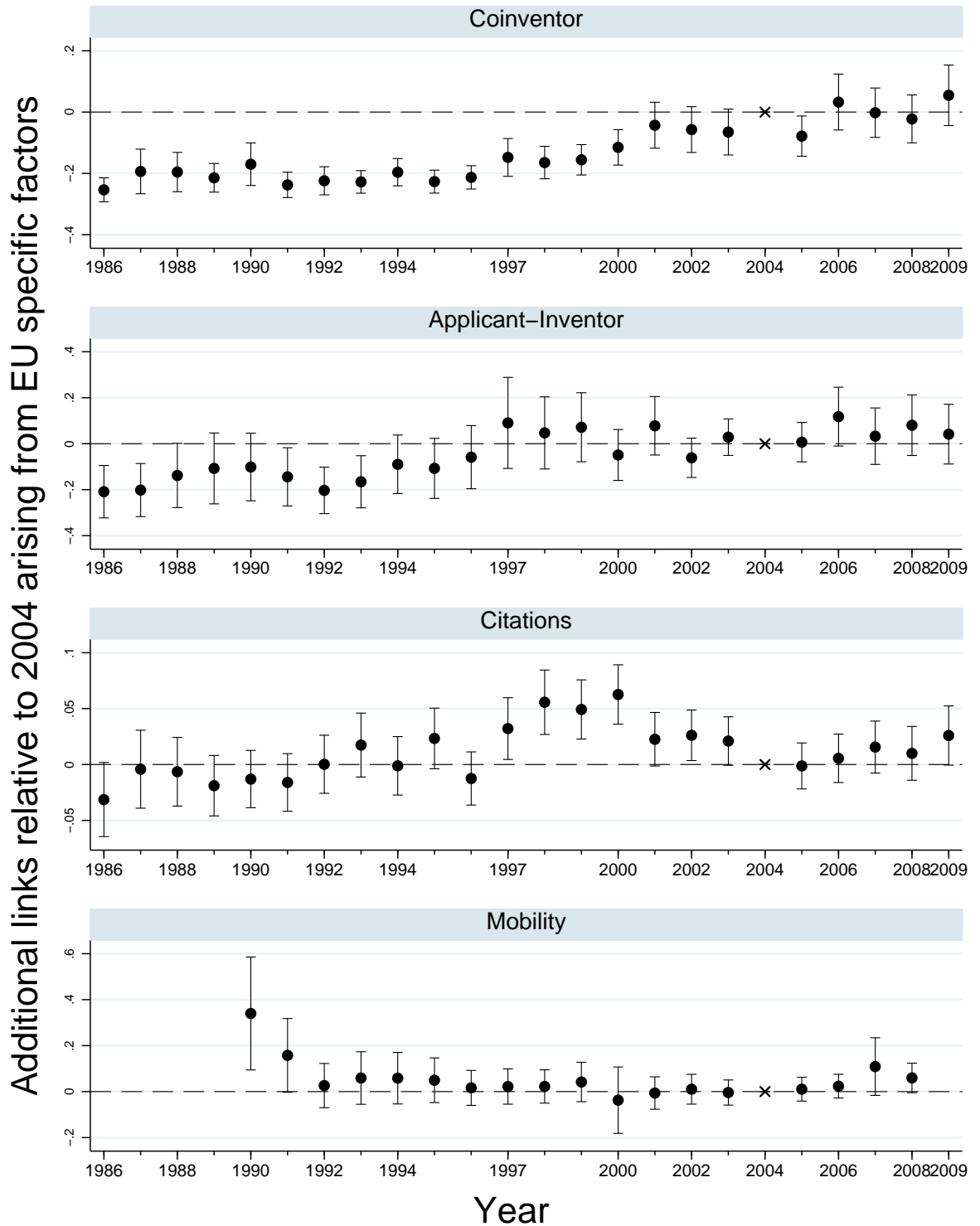
(a) (A)   **(b)** (B)

Figure 6

Regional average yearly increase in relative integration. Panel (a), intra-EU integration. Panel (b), extra-European integration.

Figure 7
EVOLUTION OF EUROPEAN INTEGRATION

Notes: Estimates derive from four separate ZINB-DiDiD-gravity models for the count of links between NUTS3 regions. In the graph we report point estimates of the average treatment effects relative to 2004 and 95% confidence intervals. The y axis reports the additional number of cross-border links for an average pair of regions (i) relative to within-border links, (ii) due to EU-specific factors as compared with non-EU OECD countries, and (iii) relative to 2004 baseline year. Marginal effects are computed for an average pair of NUTS3 regions belonging to two different EU countries in 2004. We report in Table 3 in the Appendix regression estimates from which yearly treatment effects are calculated. In the case of inventor mobility the number of non-zero link counts was too low to be modeled using ZINB, thus estimation is carried out aggregating the network at NUTS2 level.

Figure 8

DIFFERENCE BETWEEN THE HERFINDAHL INDEX OF TECHNOLOGICAL CONCENTRATION FOR TOP50 REGIONS AND ALL OTHER REGIONS, EU VS. NON-EUROPEAN REGIONS
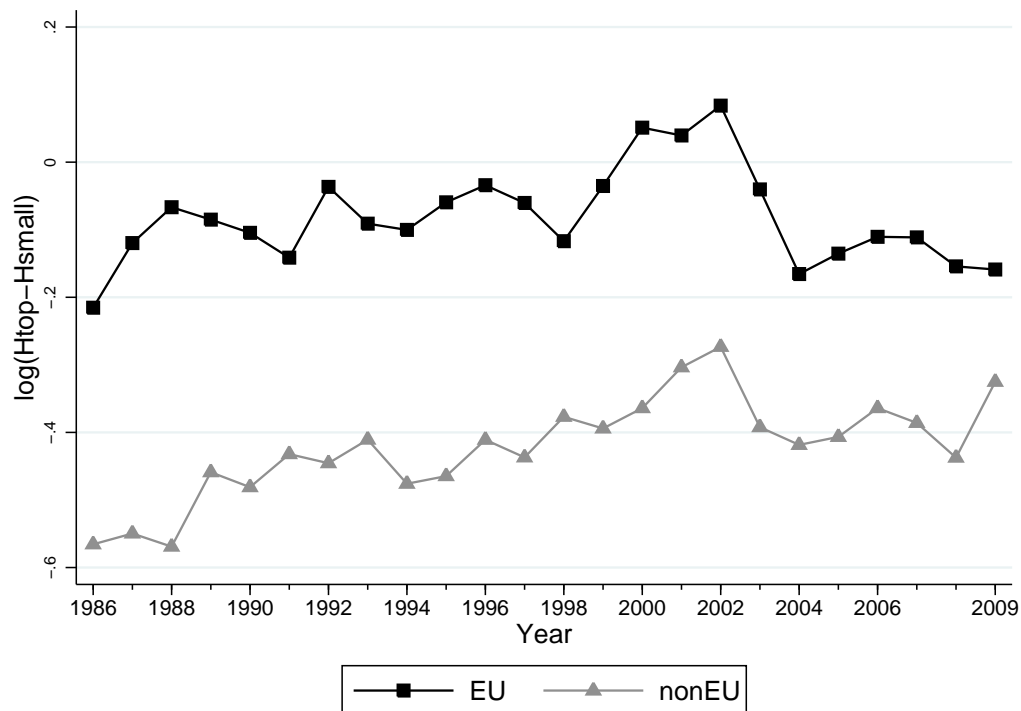
Notes: The Herfindahl index has been computed as the sum of the square of the shares of IPC classes (3rd digit) in regional patents, by year. Regions have been assigned based on the location of the inventors.

Figure 9

FLOWS BETWEEN TOP REGIONS AND SMALL REGIONS — EVOLUTION OF THE CROSS-BORDER SHARES
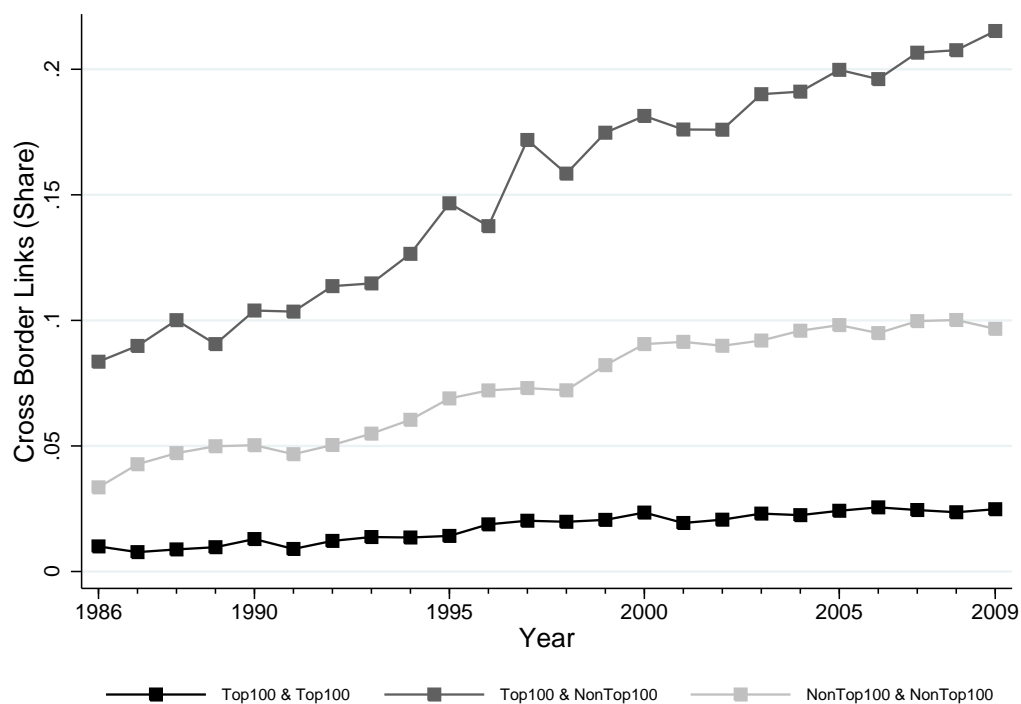
Notes: Top regions are selected counting the total number of patents over the period 1986-2009. Regional links are grouped in links of (1) both top regions, (2) one top and one non-top region, (3) both non-top regions.

Appendix

Table 1

Distance effect — Regression estimates

$y \equiv N_{(m,n)}$	Coinventor		Applicant-Inventor		Citations		Mobility	
	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI
area (base = nonEU ↮ nonEU)								
area₂ (EU ↮ EU)	-0.0501	0.4398**	-0.0218	0.2447**	0.3926**	0.6922**	0.4028**	1.1543**
	(0.0350)	(0.0313)	(0.0516)	(0.0353)	(0.0191)	(0.0162)	(0.0345)	(0.0408)
area₃ (EU ↮ nonEU)	0.2130**	-0.2868**	0.3384**	-0.3628**	0.0905**	-0.0638**	0.7929**	-0.3476**
	(0.0394)	(0.0314)	(0.0717)	(0.0547)	(0.0164)	(0.0149)	(0.0947)	(0.1249)
border	-0.9220**	1.9549**	-1.1117**	1.9128**	-0.3436**	0.7242**	-1.3374**	1.6427**
	0.0319	0.0273	0.0627	0.0429	0.0153	0.0132	0.1376	0.1764
year dummies	✓	✓	✓	✓	✓	✓	✓	✓
year_distance	✓	✓	✓	✓	✓	✓	✓	✓
distance	-0.0001**	0.0001**	-0.0001**	0.0000**	0.0000	0.0000**	-0.0002**	-0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
techdist	-2.0163**	0.8756**	-1.7550**	1.0735**	-1.6014**	0.4708**	-0.9833**	0.0458
	(0.0536)	(0.0417)	(0.0767)	(0.0539)	(0.0292)	(0.0270)	(0.0606)	(0.0700)
distance_techdist	0.0001**	0.0001**	0.0000*	0.0000	0.0000**	0.0000**	-0.0001**	-0.0002**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
neighbour	2.4128**	-26.8920**	2.0788**	-26.9908**	1.6756**	-4.0495**	1.0503**	-2.4044**
	(0.0387)	(0.5206)	(0.0524)	(0.1068)	(0.0421)	(0.2340)	(0.0426)	(0.0814)
$Size_m$	0.0007**	-0.0019**	0.0002**	-0.0058**	0.0001**	-0.0023**	0.0028**	-0.0401**
	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0008)
$Size_n$	0.0001**	-0.0003**	0.0009**	-0.0001**	0.0001**	-0.0023**	0.0022**	-0.0438**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0009)
obs	6,415,200		6,415,200		6,415,200		6,403,320	
$ln(\alpha)$	1.9004**		1.8684**		1.3393**		0.1199*	
Vuong test ~ $N(0,1)$	125.53**		157.40**		322.19**		132.19**	
l-ratio test ~ χ^2	6,758,968.3**		6,024,960.4**		8,624,418.7**		197,858.4**	

Notes: * significant at 5%; ** significant at 1%. Robust standard errors in brackets. The table reports estimates of four separate ZINB-gravity models for the count of links between NUTS3 regions (see Figure 2 for year-elasticities of the distance effect). Cross-sections of region pairs are pooled over years and estimation is carried out on the whole sample clustering standard errors at region pairs. $Size_m$ refers to the smaller of the two regions for co-inventor, while refers, respectively, to citing, to applicant's and to exit region for citations, applicant-inventor and inventor mobility. Vuong test statistics support the choice of the ZINB over a pure version NB2 ($\psi_i = 0$, $\forall i$) (Vuong, 1989; Long & Freese, 2006) and likelihood ratio tests support the choice of ZINB versus the ZIP (Long & Freese, 2006).

Table 2
Border effect — Regression estimates

$y_i \equiv N_{(m,n)}$	Coinventor		Applicant-Inventor		Citations		Mobility	
	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI
eu	0.0064	-0.2375**	-0.1147*	-0.2873**	-0.2222**	-0.3946**	-0.4385**	-0.8883**
	(0.0389)	(0.0352)	(0.0459)	(0.0341)	(0.0202)	(0.0166)	(0.0404)	(0.0555)
border	-0.9709**	2.3389**	-1.7334**	1.8759**	-0.7855**	0.6346**	-2.8596**	-0.3682
	(0.1459)	(0.0919)	(0.1009)	(0.0838)	(0.0324)	(0.035)	(0.9286)	(1.2049)
year dummies	✓		✓		✓		✓	
year_distance	✓		✓		✓		✓	
year_border	✓		✓		✓		✓	
distance	-0.0002**	0.0001**	-0.0002**	0.0001**	-0.0000**	0.0000**	-0.0001	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
techdist	-3.2972**	1.1633**	-2.9934**	1.7652**	-2.3333**	0.7581**	-1.3387**	0.2894**
	(0.0484)	(0.0409)	(0.0610)	(0.0477)	(0.0423)	(0.0334)	(0.0641)	(0.0697)
distance_techdist	0.0002**	0.0000**	0.0001**	-0.0001**	0.0001**	0.0000	-0.0001**	-0.0002**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$Size_m$	0.0008**	-0.0021**	0.0002**	-0.0065**	0.0001**	-0.0023**	0.0027**	-0.0386**
	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0008)
$Size_n$	0.0000**	-0.0004**	0.0009**	-0.0001**	0.0001**	-0.0024**	0.0022**	-0.0422**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0010)
obs	6,415,200		6,415,200		6,415,200		6,403,320	
$ln(\alpha)$	2.2959**		2.1266**		1.4315**		0.4171**	
Vuong test $\sim N(0,1)$	113.60**		149.25**		304.04**		119.97**	
l-ratio test $\sim \chi^2$	9,339,610.5**		7,031,177.5**		9,680,553.9**		201,063.2**	

Notes: * significant at 5%; ** significant at 1%. Robust standard errors in brackets. The table reports estimates of four separate ZINB-gravity models for the count of links between NUTS3 regions (see Figure 5 for year-elasticities of the distance effect). Cross-sections of region pairs are pooled over years and estimation is carried out on the whole sample clustering standard errors at region pairs. $Size_m$ refers to the smaller of the two regions for co-inventor, while refers, respectively, to citing, to applicant's and to exit region for citations, applicant-inventor and inventor mobility. Vuong test statistics support the choice of the two regions for co-inventor, while refers, respectively, to applicant's and to exit region mobility. Vuong test statistics support the choice of the ZINB over a pure version NB2 ($\psi_i = 0$, $\forall i$) (Vuong, 1989; Long & Freese, 2006) and likelihood ratio tests support the choice of ZINB versus the ZIP (Long & Freese, 2006).

Table 3
BORDER EFFECT EU vs. NONEU — REGRESSION ESTIMATES

$y_i \equiv N_{(m,n)}$	Coinventor		Applicant-Inventor		Citations		Mobility	
	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI	$y\,(\geq 0)$	ZI
border	0.7176**	-0.0607	1.7120**	1.4961**	0.0901	-0.2019**	-0.9399**	-20.3427**
	(0.1355)	(0.2968)	(0.1623)	(0.1618)	(0.0500)	(0.0741)	(0.1133)	(1.8511)
eu	0.1006**	-0.0291	0.2268**	-0.1014	-0.3748**	-0.5394**	0.1091	-0.2752
	(0.0313)	(0.0764)	(0.0442)	(0.0763)	(0.0564)	(0.0585)	(0.0732)	(0.3600)
border_eu	-2.3355**	-0.5995	-3.0723**	-0.8334**	-0.9464**	0.4161**	-1.1694**	19.4609**
	(0.1361)	(0.4006)	(0.1657)	(0.1982)	(0.0636)	(0.0904)	(0.1447)	(2.8939)
year dummies	✓		✓		✓		✓	
border-year dummies	✓		✓		✓		✓	
eu_year dummies	✓		✓		✓		✓	
border_eu_year dummies	✓		✓		✓		✓	
distance	-0.0003**	0.0001**	-0.0003**	0.0001**	-0.0001**	0.0001**	-0.0001**	0.0004**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
$Size_m$	0.0029**	-0.0363**	0.0003**	-0.0100**	0.0002**	-0.0038**	0.0015**	-0.0283**
	(0.0000)	(0.0020)	(0.0000)	(0.0006)	(0.0000)	(0.0001)	(0.0001)	(0.0051)
$Size_n$	0.0001**	-0.0001**	0.0020**	0.0002**	0.0001**	-0.0018**	0.0020**	-0.0220**
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0002)	(0.0047)
obs	3,847,704		3,151,440		7,266,600		127,015	
$ln(\alpha)$	2.1240**		2.2804**		2.2226**		0.3010**	
Vuong test $\sim N(0,1)$	152.34**		172.13**		106.8**		33.9**	
l-ratio test $\sim \chi^2$	21,128,342**		12,890,196**		12,890,196**		62,773.76**	

Notes: * significant at 5%; ** significant at 1%. Robust standard errors in brackets. The table reports estimates of four separate ZINB-DiDiD-gravity models for the count of links between NUTS3 regions (see Figure 7 for yearly treatment effect estimates). Cross-sections of region pairs are pooled over years and estimation is carried out on the whole sample clustering standard errors at region pairs. $Size_m$ refers to the smaller of the two regions for co-inventor, while refers, respectively, to citing, to applicant's and to exit region for citations, applicant-inventor and inventor mobility. In the case of inventor mobility the number of non-zero link counts was too low to be modeled using ZINB, thus estimation is carried out aggregating the network at NUTS2 level. Vuong test statistics support the choice of the ZINB over a pure version NB2 ($\psi_i = 0$, $\forall i$) (Vuong, 1989; Long & Freese, 2006) and likelihood ratio tests support the choice of ZINB versus the ZIP (Long & Freese, 2006).

Institutions
Markets
Technologies

IMT

INSTITUTE FOR ADVANCED STUDIES LUCCA