



Centre for Research &
Analysis of Migration

Discussion Paper Series

CPD 25/14

- ▶ **Inventor Diasporas and the Internationalization of Technology**
- ▶ Ernest Miguélez

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

www.cream-migration.org

INVENTOR DIASPORAS AND THE INTERNATIONALIZATION OF TECHNOLOGY

Ernest Miguélez

GREThA - UMR CNRS 5113, Université Montesquieu - Bordeaux IV¹,

Avenue Léon Duguit, 33608 Pessac – France

Tel.: +33(0)556846270

& research fellow at AQR-IREA

& research affiliate at CReAM-UCL

E-mail: ernest.miguelez@u-bordeaux.fr

Abstract

This paper documents the influence of diaspora networks of highly-skilled individuals – i.e., inventors – on international technological collaborations. Using gravity models, it studies the determinants of the internationalization of inventive activity between a group of industrialized countries and a sample of developing and emerging economies. The paper examines the influence exerted by skilled diasporas in fostering cross-country co-inventorship as well as R&D offshoring. The study finds a strong and robust relationship between inventor diasporas and different forms of international co-patenting. However, the effect decreases with the level of formality of the interactions. Interestingly, some of the most successful diasporas recently documented – namely, Chinese and Indian ones – do not govern the results.

Key words: inventors, diaspora networks, international collaborations, R&D offshoring, PCT patents

JEL: C8, J61, O31, O33, R0

¹ I thank Michel Beine, Gaetan de Rassenfosse, Stuart Graham, William Kerr, Francesco Lissoni, Catalina Martinez, Çağlar Ozgen, Hillel Rapoport, Massimo Riccaboni, Valerio Sterzi, seminar participants at GREThA-Bordeaux University (December 2013), participants in the workshop "The Output of R&D and Innovative activities: Harnessing the Power of Patent Data" (JRC-IPTS Seville, September 2013), the 7th MEIDE conference (Santiago de Chile, November 2013) and the PSDM 2013 conference (Rio de Janeiro, November 2013) for valuable comments; and Julio Raffo for helpful discussions on previous versions of this paper. However, any mistake or omission remains my own.

1. Introduction

International innovation networks are critical stepping stones for accessing frontier knowledge from the most industrialized economies, both in the form of formal information exchanges as well as through knowledge spillovers (Hall, 2011). Firms in developed countries internationalize their innovation activity in order to adapt their products to foreign production processes and foreign markets (Patel and Vega, 1999), monitor new technology developments (Guellec and van Pottelsberghe de la Potterie, 2001), and exploit technological advantages of foreign countries.² Furthermore, the complexity of technical innovation leads scientists and engineers to look for the most suitable co-inventors worldwide (Katz and Martin, 1997). To these aims, advances in modern communication technologies have contributed greatly to the overcoming of geographical obstacles to cross-country economic interactions. However, technological collaborations are, still today, primarily a national phenomenon. More than 20 years ago, Patel and Pavitt (1991) observed that the production of technology “remains far from globalized”, contrary to other features such as trade or Foreign Direct Investment (FDI). Guellec and van Pottelsberghe de la Potterie (2001) report that only 4.7% of EPO patents and 6.2% of USPTO patents in 1995 have at least one foreign co-inventor. Picci (2010) estimates this figure to be around 8% for European patents in 2005. The data used in the present paper confirm this extreme: out of all Patent Cooperation Treaty (PCT) applications listing at least 2 inventors, only 8-9% of them include inventors resident in a minimum of two countries during the 2000s. As for trade, informal barriers to cross-country economic interactions essentially explain these low rates of internationalization (Helliwell, 1998; McCallum, 1995). Transnational social networks, such as skilled migrant networks, may overcome these barriers and foster the internationalization of inventive activity. The aim of this paper is to examine this relationship.

People in the same country share historical background, cultural roots, and language. This facilitates the formation of trust and mutual understanding, eases the screening of potential partners, helps the managing and administration of a common project, and smooths the monitoring of partners' fulfillments. These factors are serious barriers to

² Examples of such location-specific technological advantages are, e.g., benefiting from a particular science base – including scientists and engineers, university research being particularly strong in certain areas, or learning from local competitors (Hall, 2011).

the internationalization of inventive activity, and largely explain the figures reported above. Other important aspects are differences in legal frameworks and the rule of law, especially regarding the issue of intellectual property rights (Foray, 1995; Montobbio and Sterzi, 2013).

Migrant networks may smooth obstacles to the internationalization of inventive activity. They create trust across national boundaries, provide information on market opportunities and, in general, reduce the transaction costs of economic interactions between countries. Diaspora networks have been studied in the context of trade (Gould, 1994), FDI (Javorcik et al., 2011; Kugler and Rapoport, 2007), international diffusion of ideas (Agrawal et al., 2011; Kerr, 2008) and further migration flows (Beine et al., 2014, 2011). In parallel, numerous papers have investigated the internationalization of R&D activities (Guellec and van Pottelsberghe de la Potterie, 2001; Patel and Vega, 1999; Picci, 2010). To the best of my knowledge, however, no study has looked at the role of highly-skilled migration in fostering international technological collaborations.³

The present paper looks at the specific issue of transnational inventive activity between developed and developing countries, which still remains an unexplored topic (see, recently, Montobbio and Sterzi, 2013). The focus of the paper is on whether a link exists between developing country inventor diasporas residing in high income economies and the opportunities for technology collaboration between the former and the latter. Hence, this subject matter is critical from a development policy perspective. The paper also aims to see whether differences emerge across the type of linkages created – co-inventorship vs. R&D offshoring networks, and, finally, the extent to which countries' characteristics govern these potential relations – whether the least similar countries, for which informal barriers are more acute, have the greatest potential to benefit from diaspora networks.

³ The most related work I found to the present paper is that by Foley and Kerr (2013), who study how the ethnic composition of US technological firms influences the internationalization of their R&D and inventive activities. I extend their work by looking at diaspora networks in several industrialized countries, not only the US. I also study diasporas coming from a wide range of origin countries, and not just nine ethnicities. Furthermore, my analysis at the country level enables me to capture broader effects of diaspora networks on international collaborations, beyond firms' responses to their share of immigrant employees.

In addition, this study extends the existing literature in a critical way. A large majority of diaspora and migration studies use total immigration data or tertiary educated immigration data retrieved from decennial censuses, which however implies a number of limitations. Thus, I use a novel dataset of inventors with migratory background as a proxy for a highly-skilled diaspora (Miguelez and Fink, 2013). Inventors make up a specific class of workers at the upper end of the skills distribution and are arguably a more homogeneous group of employees compared to the tertiary educated labor force as a whole. In addition, as the original data source comes from patent data, I am able to exploit a longitudinal dataset – 20 years, including a large number of sending and receiving countries.

To anticipate the results to come, I find a robust and sizeable effect of highly-skilled diasporas on the internationalization of inventive activity between developed, receiving countries and developing, sending economies. The effect is statistically and economically significant: a 10% increase in the inventor diaspora abroad is associated with a 2.0-2.2% increase in international patent collaborations. The evidence found survives the inclusion of a large number of controls, fixed-effects (FE), robustness checks, and identification issues. Moreover, the effect is stronger for inventor-to-inventor collaborations – co-inventorship – than for applicant-to-inventor co-patents – R&D offshoring, suggesting that diaspora effects specifically mediate interpersonal relations between co-workers.

The outline of the paper is as follows: section 2 reviews previous theoretical and empirical contributions on the relationship between migration and other international economic interactions. Section 3 presents the novel dataset on inventor migration flows and develops the methodological setting, including all the econometric concerns. Section 4 presents the results and section 5 concludes.

2. Related literature and theoretical background

Under standard trade models, trade in goods and trade in factors (such as labor) are likely to be substitutes for one another. The free movement of factors equalizes their prices and consequently commodity prices also equalize, reducing incentives to trade (Egger et al., 2012). In a similar vein, FDI flows to where labor is relatively abundant. If

migration reduces the human capital endowments of the countries of origin, migration and FDI flows can be seen as substitute ways of matching employers and employees across different countries (Kugler and Rapoport, 2007, 2005). In the Offshoring Research Network (ORN) survey conducted in 2004-2006 tracking offshoring activities of more than 1,600 European and US firms, access to qualified personnel was mentioned as the second most important offshoring factor (Manning et al., 2008). However, this global talent can be obtained in two ways: establishing operations abroad and hiring local talent; or importing talent to the firms' home countries, making offshoring and skilled immigration substitute means to hiring highly-skilled workers. Evidence in this direction is provided by Ottaviano et al. (2013), who find that in the case of the US, an increase in the ease of immigration critically reduces offshoring. The same logic applies to international inventorship teams and R&D offshoring. If firms in developed countries internationalize their innovation activities seeking for foreign pools of specialized, highly-skilled human capital, skilled immigrants in the host countries may reduce the need to locate their R&D labs abroad. At the same time, increasingly high-tech parts of the production processes of firms are being located offshore in low and middle income countries – the Indian city of Bangalore, and its development of software services, is a clear example of this – reducing incentives to migrate in the first place.

However, expatriate Indians draw distant places just like Bangalore closer to world markets, as the overseas Chinese did for Shanghai and Guangzhou previously (World Bank, 2009). A burgeoning body of literature illustrates how diaspora networks boost international economic transactions. In the migration literature, diasporas have been defined as “part of a people, dispersed in one or more countries other than its homeland, that maintains a feeling of transnational community among a people and its homeland” (Chander, 2001). Potential benefits can be realized exploiting this feeling to the advantage of the home countries, through the individuals' embedded knowledge as well as through their accessible resources – such as capital or the expatriates' network of colleagues and acquaintances.

Diasporas affect their home countries both directly and indirectly (Kapur and McHale, 2005). The direct effect is linked to the diaspora members' willingness to interact individually with their home countries, in the form of remittances, investments, or

sharing ideas and information. The indirect effects refer to the role of diaspora members in leveraging their home countries' reputation in international business networks; facilitating searching and matching between partners, customers-suppliers or in the labor market; and finally, in ensuring the contract fulfillments of the two parties involved (op. cit.). Because of their familiarity with local market needs, diasporas provide information about business opportunities in their homelands, and thus are critical in providing access to relevant information otherwise inaccessible because of cultural, language, institutional, administrative, or geographical barriers. Thus, migrant networks lower transaction costs associated with problems of incomplete information. This is particularly the case when informational difficulties are large, when involving countries with very different social and cultural backgrounds.

They also lower the transaction costs associated with the existence of asymmetric information. As Rauch (2003, 2001) posits, social networks operating across national borders, build up, or substitute for trust when contract enforcement is weak or non-existent. Indeed, diasporas create trust by establishing a kind of "moral community", which is used to transmit information about past opportunistic behavior in international business relations. The enforcement mechanism is particularly important in the absence of effective protection of contractual/property rights, and should therefore be critical in the relationships of developed-developing countries.

In the trade context, Gould (1994) finds that the stock of migrants in the United States (US) from 47 US trading partners increases US trade with these countries. This is confirmed by Rauch and Trindade (2002) and Head and Ries (1998), who find that a 10% increase in the number of immigrants increases exports by one percent and imports by three percent. Several refinements of these studies have critically shown that immigrant networks have less effect on the trade of more homogeneous products – for which prices convey the relevant information – than heterogeneous products – for which non-disclosed information is more relevant (Aleksynska and Peri, 2013). Similar conclusions emerge in the case of FDI. Javorcik et al. (2011) investigate the link between the presence of migrants in the US and US FDI to the migrants' countries of origin. They find that US FDI to sending countries is positively correlated with the diaspora of that country in the US – especially migrants with college degree qualifications (see also Kugler and Rapoport, 2007).

To the best of my knowledge, there is very little in this literature on diaspora externalities and international technology cooperation. Case studies and anecdotal evidence seem to suggest how important migrant networks for international cooperation are. Saxenian (1999) argues that skilled immigrants in the US are playing a growing role in linking domestic technology businesses to their countries of origin. Her study on Chinese and Indian immigrant engineers in Silicon Valley shows that these immigrants are uniquely placed to locate foreign partners quickly and manage complex business networks across cultural, institutional and linguistic boundaries, which is especially relevant in high-tech industries. The resulting transnational networks are likely to enhance economic opportunities both for California and for emerging regions in Asia (Saxenian, 1999).

A parallel research stream has documented the link between skilled migration and scientific collaborations. For instance, Regets (2001) finds a strong positive association between the number of foreign students awarded PhDs in the US and the degree to which scientific articles authored in their sending countries include a US author. More recently, Scellato et al. (2012) report a positive link between mobility and international research networks, for a group of surveyed scientists from 16 countries. Their study finds that around 40% of foreign-born researchers in these countries maintain research links with their homeland colleagues.

For the specific case of inventors, Agrawal et al. (2011) study knowledge flows between India and the Indian diaspora in the US, identified through inventors of USPTO patents. Kerr (2008) extends this analysis to nine ethnicities: Chinese, English, European, Hispanic, Indian, Japanese, Korean, Russian, and Vietnamese. By means of citation analysis, he confirms that knowledge diffuses internationally through ethnic networks – especially with regards to the Chinese diaspora, which also has sizeable effects on home country output. Foley and Kerr (2013) find significant effects of US firms' ethnic inventors in promoting linkages between these firms and their R&D staff home countries, in the form of knowledge flows or R&D alliances. As these authors argue, ethnic inventors in host countries are particularly apposite for helping firms to capitalize in foreign opportunities and overcome barriers to the internationalization of inventive activity. Ethnic inventors usually have the expertise essential for developing products

crucial for that particular ethnicity, giving privileged access to foreign markets and business opportunities. Obviously, they possess the language skills and cultural sensitivity necessary to promote international collaborations in their host countries, while at the same time knowing how to conduct business with their homeland colleagues. They also belong to those networks that foster trust and convey information about past opportunistic behavior across national boundaries.

To conclude this review, it is worth pointing out that some scholars have argued that lessons from the most successful Asian diasporas – namely Indian and Chinese – do not directly extrapolate to other migrant communities. In a nutshell, they argue that highly-skilled emigrants do not systematically engage in business networks and knowledge transfers with their homelands, but rather that the Indian and Chinese diasporas are so famous in being the exception rather than the rule (Gibson and McKenzie, 2012). Others argue that, while related literature is extensive in the case of the largest destination country, the US, it is limited for other receiving areas (Breschi et al., 2013). In the light of these arguments, the empirical approach presented here explores the extent to which the US experience and its top providers of foreign talent govern diaspora effects on co-inventorship and R&D offshoring, or whether results can be generalized.

3. Research methods

3.1. Inventors' international migration

In large part, the surge of empirical analysis described in the previous section responds to census-based migration datasets becoming available during the last 15 years (Carrington and Detragiache, 1998; Docquier and Marfouk, 2006; Özden et al., 2011). These datasets, broken down by skills – primary, secondary and tertiary level of education – have enabled researchers to investigate empirically the role of skilled diasporas in fostering transnational interactions, such as trade or FDI.

The present analysis is based on a new dataset of inventors with migratory background, applying for PCT patent applications, between 1990 and 2010 (for a description of the dataset, see Miguelez and Fink, 2013). The use of inventor data with migratory

background comes with two main advantages, compared to existing datasets. First, patent data (together with inventor information) are registered, and so can be organized on a yearly basis, which enables time-series variations of the data to be exploited – contrary to census data, which are collected only every 10 years – and the data are available for a large number of sending and receiving countries. Second, the level of education attained may still differ markedly among tertiary educated workers. Tertiary education can include non-university tertiary degrees, undergraduate university degrees, and postgraduate and doctorate degrees, which may not be fully comparable across different countries. Inventors on the other hand constitute a specific class of highly-skilled workers which is more homogeneous than the tertiary educated workers as a whole. They are behind the production of new knowledge and innovation that encourage economic growth and well-being. In addition to all this, using inventor information from PCT applications implies that probably the most skilled inventors are captured. PCT patent applications are clearly aimed at being extended worldwide and may hence be associated with the most valuable inventions (Guellec and Van Pottelsberghe de la Potterie, 2002; Jensen et al., 2011; van Zeebroeck and van Pottelsberghe de la Potterie, 2011).⁴ Furthermore, to the best of my knowledge, PCT patent applications are the only ones recording this type of information. This is because not all countries are PCT contracting states, and only nationals or residents of a PCT contracting state can file PCT applications. In order to verify that applicants meet at least one of the two eligibility criteria, the PCT application form asks for both nationality and residence. In parallel to this, it turns out that US patent application procedures require the applicant of a patent also to be the inventor. If a given PCT application includes the US as a country in which the applicant has considered pursuing a patent – a so-called designated state in the application – all inventors are listed as applicants and their residence and nationality information are, in principle, available – in fact, this is the case for the majority of applications.

⁴ The use of patent data does not come without limitations . Apart from the well-known issues of varying patent quality and the fact that not all inventions are patented, more problematic for the present analysis is that the observation of both migration and collaborations is based on successful outcomes (the patent application). Thus, potential biases created by the impossibility to observe migration and collaborations without a successful output are addressed by means of the instrumentalization strategy described below.

All in all, between 1990 and 2010, the share of inventors' records for which we can retrieve nationality and residence information is pretty high, around 80% of the cases.⁵ Admittedly, this coverage is unevenly distributed over time – around 60-70% during the 1990s and 70-95% during the 2000s, as well as across countries – US (66%), Canada (81%), the Netherlands (74%), Germany (95%), the United Kingdom – UK (92%), France (94%), Switzerland (93%), China (92%) and India (90%), among others.⁶ Once individual-level data are retrieved, I aggregate across pairs of countries and years.⁷ In particular, I treat each record in the patent database as if it were a different individual and compute diaspora variables for annually repeated time-windows of five years.

Out of all records with complete information – about 5 million, around 9-10% have migratory background – i.e., residence different from nationality. Figure 1 depicts the evolution of the share of inventors with migratory background (solid line), alongside the same figures broken down by a number of selected receiving countries/areas. As can be observed, the share of worldwide migrant inventors has steadily increased over time.⁸ Among the most receiving countries of the world, Canada, Australia and, notably, the US, stand out as being the primary receiving countries, when compared to their resident stock of inventors. On the other hand, Japan is, and has been over the years, one of the developed countries with a smaller share of inventor immigrant population.

[Figure 1 about here]

Meanwhile, technology-leading European countries, such as Germany or France, lag way behind compared to the US (Figure 2). On the contrary, smaller European countries rank even better than the US in terms of immigration rates of inventors – notably, Switzerland, Ireland or Belgium. However, the exceptional performance of the US in

⁵ The use of the word “record” here signifies the unique combination of “inventor name” and “application number”.

⁶ To address this inconsistent coverage of migration information over time, I repeated the analysis splitting the sample into shorter time windows. No important differences arise regarding the main conclusions of the study.

⁷ I use the priority date of applications to allocate individuals in time. By “priority date” I mean the first year the patent was applied worldwide.

⁸ In order to make these figures comparable, it is worth looking at differences with other migration datasets. While 8.62% of inventors of PCT patents have a migratory background in 2000, data compiled by Docquier and Marfouk (2006) or Beine et al. (2007) show that general migration rates in 2000 for populations aged 25 years and over were estimated around 1.8%, including 1.1% of immigrants among the unskilled population, 1.8% among populations with secondary education, and 5.4% among populations with tertiary education.

attracting talent is notorious when considering only immigrant inventors coming from low and middle income economies, as can also be seen in Figure 2.

[Figure 2 about here]

Table 1 shows the top-20 most populated corridors (left panel). As expected, the US stands out as the most typical choice for destination country, while most origins are other high income economies. The nameable exceptions are the top two corridors – China-US and India-US – with middle income country origins. Other middle income economies are also important sources of inventors during the period 2001-2010 e.g., Russia, Turkey, Iran, Romania and Mexico (right panel).

Admittedly, the bulk of inventor migration is concentrated in North-North corridors, although South-North corridors, which are the object of analysis in this paper, are also sizeable. In fact, the South-North corridor has gained prominence over the North-North one, mostly due to the massive migration flows of Indian and Chinese inventors to the US (from around 20% at the beginning of the 1990s, to almost 40% at the end of the 2000s). A similar pattern emerges when looking at cross-country PCT co-patenting figures at the inventor level, e.g. co-inventorship: PCT co-inventions listing at least one non-OECD inventor have risen exponentially from 14% at the beginning of the 1990s to around 40% in 2012. The lower costs of conducting R&D activities in developing economies largely explain this surge (Montobbio and Sterzi, 2013). Moreover, due to recent experiences of economic growth and technological development of some emerging economies, access to qualified personnel and location-specific knowledge pools have driven co-inventorship and R&D offshoring with these countries (Thursby and Thursby, 2006).

[Table 1 about here]

3.2. Empirical approach

The gravity model to be estimated takes the following form:

$$COPAT_{ijt} = e^{\beta_0} \cdot DIASPORA_{ijt}^{\beta_1} \cdot Z_{ijt}^{\gamma_n} \cdot e^{\tau_i} \cdot e^{\tau_j} \cdot e^{\delta_t} \cdot \varepsilon_{ijt} \quad (1)$$

where $COPAT_{ijt}$ stands for the number of collaborations between i's developing country (out of 67) and j's developed country (out of 20), for year t.⁹ β_1 is the parameter of interest in this work, while $DIASPORA_{ijt}$ is the focal variable and is computed in two main ways. First, the number of inventor nationals of country i residing in country j, for annually repeated 5-year time-windows; second, the share of inventor nationals of country i residing in country j out of all inventors residing in country j, for annually repeated 5-year time-windows. Z_{ijt} is a set of bilateral and attribute control variables, and τ_i , τ_j , and δ_t are, respectively, developing, developed and time FE. ε_{ijt} denotes the error term.

Log-linearizing equation (1) and using OLS techniques would be a straightforward estimation method. However, cross-country co-patents are rare phenomena, which translate into a dependent variable with a very large proportion of zeros, making the logarithmic transformation of these observations impossible. Dropping these zero observations or adding an arbitrary constant to enable logarithmic transformation would be clearly misleading (Burger et al., 2009). In addition, Santos Silva and Tenreyro (2006) show that log-linearizing equation (1) may induce a form of heteroskedasticity of the error term because of log-transformation of the data, making OLS estimations inconsistent. Instead, the authors suggest estimating the multiplicative form of the model using Poisson pseudo-maximum likelihood (PPML), which also provides a natural way of dealing with zero co-patenting and the extreme skewness of the dependent variable, intrinsically heteroskedastic with variance increasing with the mean (Cameron and Trivedi, 1998). Thus, I estimate equation (1) by means of PPML using the fact that the conditional expectation of $COPAT_{ijt}$ in (1) can be written as the following exponential function

$$E(COPAT_{ijt} | X_{ijt}) = \exp[\beta_0 + \beta_1 \ln DIASPORA_{ijt} + \gamma_n \ln Z_{ijt} + \tau_i + \tau_j + \delta_t + \varepsilon_{ijt}]. \quad (2)$$

⁹ Appendix 1 lists the countries used in this study. Note that I included some high income countries among the list of developing economies (e.g. Luxemburg). Removing them from this group does not alter the results.

3.3. Data

Dependent variable

International co-patent data are retrieved from PCT applications (WIPO IPSTATS databases). I first focus on co-patenting at inventor level – co-inventorship. I add up by year all the co-inventions between inventors residing in country i and inventors residing in country j . To be precise, I include 67 developing/emerging/transition countries, on the one hand, that co-invent with 20 developed countries, on the other, where diasporas from the former countries reside. If inventors from more than two countries participate in the patent, I count an international co-inventor for each country-pair, irrespective of the total number of countries involved in that particular invention.¹⁰

Next, I also look at the role of inventor diasporas in fostering R&D offshoring to their homelands. I measure R&D offshoring using patent applications in which at least one applicant is a resident from country j (developed) and simultaneously at least one inventor is a resident from country i (developing/emerging/transition) – similar R&D offshoring measures in the context of internationalization of inventive activity are used in Guellec and van Pottelsberghe de la Potterie (2001), Harhoff et al. (2013) and Thomson et al. (2013). Again, when inventors come from various countries, I compute a single co-patent for each bilateral i - j pair.

It is worth mentioning that previous studies on the determinants of international co-patenting use information from single patent offices only, with few exceptions (Martínez and Rama, 2012; Picci, 2010). This practice when studying the internationalization of inventive activity between a large number of countries is likely to deliver biased estimates due to the ‘home bias effect’. The ‘home bias effect’ emerges when using patent data from one single office for cross-country analysis. Since patents at the USPTO, EPO, or JPO, for instance, protect innovation within their respective geographical areas, they are preferred by domestic firms, and thus their innovative capability is overestimated with respect to foreign firms. Using data from the PCT

¹⁰ Note that I computed other more complex measures of international co-invention, following Picci (2010) or Hoekman et al. (2010). No important changes arise – results are provided upon request.

mitigates this effect because these patents are by definition international and applicants from all countries are equally likely to apply through the PCT system – provided that the applicants are either nationals or residents of a PCT member state.¹¹ For this same reason, this paper provides additional added value to the literature on inventor migration, which so far has focused on the US and using data only from the USPTO (Breschi et al., 2013).¹²

Control variables

Control variables include geographical, linguistic, cultural, and historical barriers to cross-country collaborations. In particular, I include the great circle distance between the most populated cities of countries (measured in km), a dummy variable indicating whether two countries share a common border, a dummy variable valued 1 if the same language is spoken in both countries, and a dummy variable valued 1 when the two countries share the same colonial past – these variables come from the CEPII distance database (Mayer and Zignago, 2011).

Two additional variables help in controlling for cultural similarities between country pairs: first, I compute an index of language similarity. It is reasonable to expect that people whose languages share common roots will also share similar cultural backgrounds. To compute this index, I assign one single language to every country, and using information on the classification of languages provided by the Ethnologue Project, I compute a language similarity index based on the distance between branches in this classification.^{13,14} I then add together the number of branches that coincide between each pair of languages and divide the result by the sum of branches of each of the two languages (in order to take into account the fact that the granularity of branches

¹¹ Other biases inherent to the existence of multiple jurisdictions and patent offices are discussed in de Rassenfosse et al. (2013) – such as the non-random choice of patent office. Again, the use of PCT applications should mitigate these biases.

¹² For robustness, I repeated all the estimations using “Triadic Patent Families” (TPF) to build the dependent variables (OECD Triadic Patent Families database, January 2014). TPF consist of a set of patents filed at the European Patent Office (EPO), the Japan Patent Office (JPO) and granted by the United States Patent and Trademark Office (USPTO) that share one or more priority applications.

¹³ For example, the linguistic classification of Portuguese, Swedish, and Danish, from the largest, most inclusive grouping to the smallest, is: Indo-European, Italic/Romance, Italo-Western, Western, Gallo-Iberian, Ibero-Romance, West Iberian, Portuguese-Galician (Portuguese); Indo-European, Germanic, North East, Scandinavian, Danish-Swedish, Swedish (Swedish); Indo-European, Germanic, North East, Scandinavian, Danish-Swedish, Danish-Riksmal, Danish (Danish).

¹⁴ www.ethnologue.com, accessed 20th May 2014.

may not be the same across languages). As a result, I obtain an index between 0 and 1, where 0 means complete dissimilarity and 1 means that these two languages are almost the same in linguistic terms.¹⁵ Second, the religious heritage of countries is a critical element of their culture and identity (Guiso et al., 2009). Countries with similar religious roots, culturally closer, are likely to interact more (for an application in the trade literature, see, among others, De Groot et al. (2004). I proxy religion similarity with an index built as follows: for each country, I retrieve data on the percentage of population adhering to one of eight major religions (data from the CIA World Factbook dataset). I then compute the following formula for each country pair, which results in a variable ranging from 0 (no believers in common) to 1:

$$\begin{aligned}
 Religion_Sim_{ij} = & (\%muslim_i * \%muslim_j) + (\%catholic_i * \%catholic_j) + \\
 & (\%orthodox_i * \%orthodox_j) + (\%protestant_i * \%protestant_j) + \\
 & (\%hinduism_i * \%hinduism_j) + (\%buddhist_i * \%buddhist_j) + \\
 & (\%eastern_i * \%eastern_j) + (\%judaism_i * \%judaism_j)
 \end{aligned} \tag{3}$$

I also control for the intensity of economic linkages between countries using the share of bilateral trade (exports plus imports, EXP+IMP) between a given pair over their total trade (COMTRADE data). Trade is a conduit of information which may foster technological partnerships too, whilst it might be linked to the presence of migrants at the same time. I also account for the common technological specialization of country-pairs introducing an index of technological distance measured as

$$Tech.distance_{ij} = 1 - \frac{\sum f_{ih} f_{jh}}{(\sum f_{ih}^2 \sum f_{jh}^2)^{1/2}}, \tag{4}$$

where f_{ih} stands for the share of patents of one technological class h according to the IPC classification (out of 300 technological classes in the subdivision chosen) of country i, and f_{jh} for the share of patents of one technological class h of country j. Values of the index close to the unity indicate that a given pair of countries are

¹⁵ I arbitrarily set to 0 this variable when the countries share exactly the same language, in order to avoid collinearity with the variable ‘same language’.

technologically different, and values close to zero indicate that they are technologically similar (Jaffe, 1986). Again, I use PCT patents to compute this index.

Finally, two additional attribute variables of individual countries are used in order not to bias the point estimates of my focal regressors. In particular, I introduce the number of PCT patents per country, for 5-year annually repeated time-windows. This variable controls for the size of country innovation systems, which clearly determines a country's capacity to collaborate with foreigners, as well as its capacity to attract inventors from abroad or send them to other locations. In addition, I retrieve GDP per capita from the World Development Indicators (World Bank), expressed in US\$ 2005 at PPP, in order to capture the market potential of countries as well as their capacity to innovate. Appendix 2 contains summary statistics of the variables included in the models, as well as the correlation matrix.

Note that I lag all time-variant explanatory variables one period in order to lessen potential biases caused by system feedbacks. Notwithstanding this common practice, other sources of endogeneity and biased estimates are likely to arise. Hence, I discuss alternative solutions in the results section.

4. Results

4.1. Baseline estimations

Table 2 presents the results of baseline PPML estimations, with robust, country-pair clustered standard errors. I use two alternative focus explanatory variables: the size of the bilateral diaspora and the share of the bilateral diaspora over the number of inventors in receiving countries. Columns (1) and (2) regress international co-inventorship against these two focal variables separately, plus individual-country and time FE. The effect of the two variables is positive and statistically significant. Columns (3) and (4) further introduce a number of control variables. The focal variables remain statistically significant, although their point estimates are somewhat reduced. In particular, column (3) shows an elasticity of 0.20. That is to say, a 10% increase in the size of the inventor diaspora abroad is associated with a two-percent increase in international patent collaborations, which is also economically meaningful. This result

is of the same order of magnitude as estimates for the case of diasporas and trade (Head and Ries, 1998; Rauch and Trindade, 2002). The estimated coefficient for the diaspora variable as a proportion of the local inventors is of similar magnitude (in statistical terms). However, the exact interpretation of its elasticity is somewhat tricky – the variable is the log-transformation of a ratio. I will therefore focus my attention on the total diaspora coefficients hereafter.

[Table 2 about here]

The results for the remaining explanatory variables are interesting in themselves. As expected, physical distance between the most populated cities exerts a negative influence on the likelihood for cooperation across national boundaries, although sharing a common border barely affects co-inventorship. Common language has a strong positive estimated effect on collaborations between inventors of different countries. Other proxies for cultural similarity, such as language and religion proximity, exert a strong positive effect too. However, historical links between country pairs expressed by their colonial past are not significant. As expected, bilateral trade is positive and significant, while technological distance between countries, i.e. how distant are countries in their technological specialization, exerts a negative influence on bilateral co-patents. Finally, both attribute variables – total number of patents and GDP per capita – are significant for the case of origin countries, but not for destinations. Thus, it appears that differences across industrialized economies in terms of technological and economic development are relatively minor and are picked up by their country FE.

In columns (5) and (6) I look at R&D offshoring – co-patents between applicants in developed countries and inventors in developing economies. Comparing the estimates with those of columns (3) and (4), interesting results emerge. First and foremost, the estimated elasticity of inventor diaspora size is notably reduced in these latter estimations – less than a half compared to columns (3) and (4). That is, diaspora networks particularly mediate interpersonal relations between co-workers. Meanwhile, they have a more nuanced effect on transnational employer-employee linkages.

Second, geography per se does not play a significant role in explaining R&D offshoring. The diaspora and geography results put together seem to suggest that personal face-to-

face relations and trust building are critical in explaining co-inventorship – where contracts are usually more tacit and contract enforcement is difficult, but less important in explaining more formal and hierarchical relationships – such as those represented by offshoring relationships, where probably explicit, written contracts are the rule.

Other remarkable differences are worth reporting. For instance, the coefficient associated with colonial past increases its point estimate and now becomes significant. That is, historical ties between the former metropolis and its formal colonies seem to have left an enduring effect over time that, still today, influences innovation networks across national borders. Finally, the common specialization of countries seems to play a greater role too when looking at applicant-to-inventor co-patents, as compared to inventor-to-inventor collaborations.

Furthermore, I have used these specifications (columns (3) and (5)) to include the focal variable interacted with time dummies, in order to explore the effects of skilled diasporas over time. Figures 3 and 4 present the estimated coefficients of the interaction variables, and show quite interesting, although expected, results. As can be seen, diaspora effects over time on co-inventorship and R&D offshoring present a marked decreasing trend. Quite likely, the use of information and communication technologies (ICT) has made the utilization of international ethnic ties less relevant now than 20 years ago. Similarly, the economic and institutional development of emerging economies has also contributed to mitigate the role of skilled diasporas in reducing the costs of asymmetric information – e.g., think of the role of intellectual property in these emerging countries in the aftermath of their subscription to TRIPs, the Trade Related Intellectual Property Agreements that comes with membership of the World Trade Organization (WTO).

[Figures 3 and 4 about here]

Table 3 mimics estimations (3) through (6) of Table 2 but controlling for time-variant multilateral resistance. While country FE control for average multilateral resistance to collaborate over time (Feenstra, 2004), some elements of this multilateral resistance are likely to be time-variant and may not be picked up by the attribute variables included

(Adam and Cobham, 2007).¹⁶ In consequence, Table 3 includes country FE plus country-specific time dummies, and repeats the main estimations, focusing attention only on bilateral variables. Some nameable differences with respect to Table 2 emerge, like a reduced role of distance in explaining inventor-to-inventor collaborations. However, the focal variables remain positive and strongly significant, and they present coefficients that are slightly larger than previously.

[Table 3 about here]

4.2. Identification: cultural proximity and instrumental variables

Table 4 comes back to baseline specifications and adds interaction terms between the inventor diaspora variable and different dimensions of cultural proximity between countries – common language, language similarity, common colonial past, and religion similarity. Given that transnational migrant networks mitigate the costs of incomplete information beyond country boundaries, one would expect their impact to be stronger for country pairs exhibiting larger informational frictions. Hence, negative and significant interaction terms will provide evidence on the least similar countries relying more on diaspora externalities than pairs of countries that are culturally closer. Results (Table 4) partially confirm this extreme: all the interaction terms included are negative. Admittedly, though, only the interactions with language similarity and colonial past are statistically significant, but not the others.¹⁷

As in Kugler et al. (2013), I interpret these negative coefficients as evidence of a causal link between inventor diasporas and international co-inventive activity. Indeed, a main concern of my analysis is the possibility of omitted variables driving both migration and co-patenting at the same time. However, if unobserved confounding factors remain, they should work in such a way that they are capable of explaining not only the main results – the diaspora-co-patenting relation – but also the differentiated effect of diaspora networks across different cultural dimensions, which I find unlikely.

¹⁶ For an application to the migration literature, see Bertoli and Fernández-Huertas Moraga (2013).

¹⁷ The same estimation procedure using R&D offshoring as the dependent variable delivers similar results (negative coefficients of the interactions), but not significant. This is further evidence of the critical role of diasporas for worker-to-worker collaborations, and their more nuanced effects for the case of more hierarchical, R&D offshoring relations.

[Table 4 about here]

For robustness, I also provide instrumental variables estimates and check the validity of the results. Potential and available candidates for such a role are (i) the size of the bilateral diaspora between countries i and j in the 1960s – and its square (data from Özden et al. (2011)), and (ii) the size of the unskilled diaspora original from country i residing in country j in 1990 (migrants with only primary education) (data from Docquier et al. (2009)) – and its square. First, the stocks of migrants in the 1960s are likely to affect the current stocks of highly-skilled migrants through network effects favoring further migration flows over the long run. Quite probably, they are uncorrelated with current levels of cross-country collaborations, apart from influence through current skilled diasporas. Similarly, the current stocks of migrants with only primary education probably correlate with current stocks of highly-skilled diasporas. The relation between existing diasporas and existing migration flows not only operates at a labor market level, but also among ethnic communities operating across different skills groups. Large stocks of unskilled immigrants in a given country will mean the existence of attractive factors – e.g., amenities – which are also attractive to highly-skilled immigrants (Hunt and Gauthier-Loiselle, 2008). On the other hand, uneducated migrants should play an inexistent role in boosting co-inventorship or R&D offshoring with their homelands – justifying their exclusion from the main equations, apart from their effects through inventor diasporas. Moreover, unskilled diaspora data come from the 1990 census – which accounts for the unskilled migrant flows of the 1980s – so as to be more confident that they are unaffected by unobserved factors influencing co-patenting patterns between 1990 and 2010.

Table 5 presents GMM estimations of the PPML – see Windmeijer and Silva (1997). Note that the value of the F-test statistic, 344.58, is well above 10, which is usually considered a good threshold, and so the instruments cannot be judged as weak. Moreover, Hansen J statistics for mutual consistency of available instruments are provided and they do not reject the null hypothesis that the excluded instruments are valid and uncorrelated with the error term, so there are no over-identification problems. Column (1) shows GMM estimates using co-inventorship as the dependent variable. It shows a positive and statistically significant relationship between inventor diasporas

and international co-inventorship. The GMM results are slightly stronger in terms of magnitude of estimated coefficient relative to former PPML. Thus, analysis suggests that ignoring the endogeneity issue tends to underestimate the effect of migration on international co-inventorship. Note, however, that the difference is small and therefore the coefficients are comparable. On the other hand, results for the case of R&D offshoring remain positive, but not significant (column (2)). Again, this is further evidence of the critical role of highly-skilled diasporas for worker-to-worker co-inventorship, and their less important effect for applicant-to-inventor relations.

[Table 5 about here]

4.3. Sector heterogeneity

Use of patent data allows me to identify sector-specific particularities in the relationship between international co-patenting and diaspora networks. In particular, I follow Schmoch's (2008) classification of IPC codes into 35 technology fields, and group them into 5 broad sectors – namely, electrical engineering, instruments, chemistry, mechanical engineering, and others.

Columns (1) through (5) of Table 6 split collaborations and diaspora data into the five sectors and re-estimate the baseline model on the determinants of international co-inventorship. The coefficient on the inventor diaspora is positive and statistically significant at a one-percent level in all the specifications. Moreover, and contrary to other variables, differences across sectors are not large, witnessing to the importance of networks regardless of the technology being analyzed. The lower part of Table 6 repeats the estimates for the R&D offshoring case. Again, the focal variable in this study presents positive coefficients in all sectors analyzed, although non-significant point estimates for some of them. Moreover, differences in coefficients are remarkable.

[Table 6 about here]

4.4. Are China and India different after all?

Next, I look at the robustness of my results once the main players are removed from the analysis. This is motivated by the observation that the majority of studies look at the case of the largest receiving country, i.e. the US, and its main providers of skilled talent – i.e., China, India, and other Asian economies. Thus, for instance, Agrawal et al. (2011) provide results for the case of the Indian inventor diaspora in the US. Kerr (2008) extends the analysis to only nine ethnicities – Chinese, English, European, Hispanic, Indian, Japanese, Korean, Russian, and Vietnamese – finding that only the Chinese inventor diaspora successfully diffuses knowledge back to its homeland. (Saxenian, 2006, 2002, 1999) studies Indian and Chinese migrant entrepreneurs. These and related studies generally focus only on one destination country, the US, while migration and international business networks are multi-country phenomena (Breschi et al., 2013). In the light of this, some scholars argue that lessons from case studies on China and India cannot extrapolate to other migrant communities – that is, it is difficult to say whether highly-skilled emigrants systematically engage in business networks and knowledge transfers with their homelands or rather that the Indian and Chinese diasporas are so famous for being an exception rather than the rule (Gibson and McKenzie, 2012).

In order to explore this issue, Table 7 repeats co-inventorship estimations – with and without country-specific time dummies – but removing from the sample either the BRICS countries (Brazil, Russia, India, China and South Africa), or the US, or both. Contrary to the arguments posited by Gibson and McKenzie (2012), among others, the coefficient accompanying the diaspora variable remains strongly significant and economically meaningful in all models, and barely lower when compared to previous estimates.

[Table 7 about here]

4.5. Further robustness analysis

To further check the robustness of my results, Table 8 runs the baseline specification using different estimation methods. In particular, it includes OLS, Negative Binomial, and zero-inflated Poisson and Negative Binomial. For all cases the estimated coefficients are slightly larger compared to Table 2, but they are fairly comparable. In

addition to these checks, in unreported tables I re-estimate all models using TPF patents to build the dependent variables. Because TPF applications protect inventions aimed at the international sphere through the three largest patent offices, they are supposed to not suffer the home-bias effect while, at the same time, constituting the most economically and technologically valuable inventions (de Rassenfosse and van Pottelsberghe de la Potterie, 2009). Interestingly, coefficients remain positive and strongly significant – and with comparable values.

[Table 8 about here]

5. Conclusion

This paper examines the impact of highly-skilled migrant networks in high income countries on the internationalization of inventive activity between high income and developing economies – measured as cross-country PCT co-patenting (co-inventorship and R&D offshoring). In order to study this relationship, I make use of a unique dataset on inventors with a migratory background. To my knowledge, there have been no previous attempts to measure the mentioned links, and therefore this constitutes the main contribution of the paper.

The results show a strong and positive association between highly-skilled diasporas and the internationalization of inventive activity between developed and developing countries. The effect is statistically and economically significant: a 10% increase in the inventor diaspora abroad is associated with a 2.0-2.2% increase in international patent collaborations at the level of inventors. The effect found is robust to the inclusion of a bunch of controls and FE, including individual country FE interacted with time FE. Given the variables included, the econometric approach – including instrumental variables estimates, and the robustness checks performed, I am fairly confident that my focal regressors do not pick up any confounding effect that might bias their point estimates. These findings do not suffice to conclude that a ‘brain gain’ exists that outreaches the loss of highly-skilled human capital of sending economies, although they are undeniably necessary elements.

Interestingly enough, the effect, although relatively diminished, does not depend on the remarkable performance of particular diasporas abroad, such as Chinese or Indian inventors. Equally, results are not particularly driven by the country both attracting the largest number of migrant inventors and concentrating a significant proportion of North-South international collaborations, i.e., the US. It seems therefore that inventor diaspora effects are exhausted at relatively low levels of highly-skilled diasporas (for similar results for the trade-migration relationship, see Egger et al, 2012).

The results also suggest that highly-skilled diaspora effects weaken dramatically in the case of R&D offshoring – collaborations between applicants in developed countries and inventors in developing ones. These results seem to suggest that personal face-to-face relations and trust building are critical to explain co-inventorship – where contracts are usually more tacit and contract enforcement is difficult, but are less important in explaining more formal and hierarchical relationships – where probably explicit, written contracts are the rule. I interpret these results as evidence that substitutability between skilled immigration and R&D offshoring outreaches the diaspora effect, while this is not the case at the level of the job market for inventors, where diaspora networks play a referral role. Further research, possibly at the firm and inventor levels, will shed more light on this particular issue.

References

- Adam, C., Cobham, D., 2007. Modelling multilateral trade resistance in a gravity model with exchange rate regimes (CDMA Conference Paper Series No. 0702). Centre for Dynamic Macroeconomic Analysis.
- Agrawal, A., Kapur, D., McHale, J., Oettl, A., 2011. Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. *Journal of Urban Economics* 69, 43–55. doi:10.1016/j.jue.2010.06.003
- Aleksynska, M., Peri, G., 2013. Isolating the Network Effect of Immigrants on Trade. *The World Economy* n/a–n/a. doi:10.1111/twec.12079
- Beine, M., Docquier, F., Özden, Ç., 2011. Diasporas. *Journal of Development Economics* 95, 30–41. doi:10.1016/j.jdeveco.2009.11.004
- Beine, M., Docquier, F., Rapoport, H., 2007. Measuring International Skilled Migration: A New Database Controlling for Age of Entry. *World Bank Econ Rev* 21, 249–254. doi:10.1093/wber/lhm007
- Beine, M., Noël, R., Ragot, L., 2014. Determinants of the international mobility of students. *Economics of Education Review* 41, 40–54. doi:10.1016/j.econedurev.2014.03.003
- Bertoli, S., Fernández-Huertas Moraga, J., 2013. Multilateral resistance to migration. *Journal of Development Economics* 102, 79–100. doi:10.1016/j.jdeveco.2012.12.001
- Breschi, S., Lissoni, F., Tarasconi, G., 2013. Inventor Data for Research on Migration & Innovation: A Survey and a Pilot.
- Burger, M., van Oort, F., Linders, G.-J., 2009. On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. *Spatial Economic Analysis* 4, 167–190. doi:10.1080/17421770902834327
- Cameron, A.C., Trivedi, P.K., 1998. *The analysis of count data*. Cambridge University Press, Cambridge.
- Carrington, W., Detragiache, E., 1998. *How Big Is the Brain Drain?* International Monetary Fund.
- Chander, A., 2001. Diaspora Bonds (2001). *New York University Law Review* 76, 1005–1043.
- De Groot, H.L.F., Linders, G.-J., Rietveld, P., Subramanian, U., 2004. The Institutional Determinants of Bilateral Trade Patterns. *Kyklos* 57, 103–123. doi:10.1111/j.0023-5962.2004.00245.x
- De Rassenfosse, G., Schoen, A., Wastyn, A., 2013. Selection bias in innovation studies: A simple test. *Technological Forecasting and Social Change*. doi:10.1016/j.techfore.2013.02.012
- De Rassenfosse, G., van Pottelsberghe de la Potterie, B., 2009. A policy insight into the R&D–patent relationship. *Research Policy* 38, 779–792. doi:10.1016/j.respol.2008.12.013
- Docquier, F., Lowell, B.L., Marfouk, A., 2009. A Gendered Assessment of Highly Skilled Emigration. *Population and Development Review* 35, 297–321. doi:10.1111/j.1728-4457.2009.00277.x
- Docquier, F., Marfouk, A., 2006. *International Migration by Education Attainment (1990-2000)-Release 1.1*. working paper.
- Egger, P.H., von Ehrlich, M., Nelson, D.R., 2012. Migration and Trade. *The World Economy* 35, 216–241. doi:10.1111/j.1467-9701.2011.01429.x
- Feenstra, R.C., 2004. *Advanced international trade: theory and evidence*. Princeton University Press, Princeton, N.J.

- Foley, C.F., Kerr, W.R., 2013. Ethnic Innovation and U.S. Multinational Firm Activity. *Management Science*. doi:10.1287/mnsc.1120.1684
- Foray, D., 1995. The economics of intellectual property rights and systems of innovation: the persistence of national practices versus the new global model of innovation. *Technical change and the world economy: convergence and divergence in technology strategies* 109–133.
- Gibson, J., McKenzie, D., 2012. The Economic Consequences of “Brain Drain” of the Best and Brightest: Microeconomic Evidence from Five Countries*. *The Economic Journal* 122, 339–375. doi:10.1111/j.1468-0297.2012.02498.x
- Gould, D.M., 1994. Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows. *The Review of Economics and Statistics* 76, 302–316.
- Guellec, D., van Pottelsberghe de la Potterie, B., 2001. The internationalisation of technology analysed with patent data. *Research Policy* 30, 1253–1266. doi:10.1016/S0048-7333(00)00149-9
- Guellec, D., Van Pottelsberghe de la Potterie, B., 2002. The Value of Patents and Patenting Strategies: Countries and Technology Areas Patterns. *Economics of Innovation and New Technology* 11, 133–148. doi:10.1080/10438590210896
- Guiso, L., Sapienza, P., Zingales, L., 2009. Cultural Biases in Economic Exchange? *The Quarterly Journal of Economics* 124, 1095–1131. doi:10.1162/qjec.2009.124.3.1095
- Hall, B.H., 2011. The internationalization of R&D (UNU-MERIT Working Paper Series No. 049). United Nations University, Maastricht Economic and social Research and training centre on Innovation and Technology.
- Harhoff, D., Mueller, E., Reenen, J.V., 2013. What are the Channels for Technology Sourcing? Panel Data Evidence from German Companies (CEP Discussion Paper No. dp1193). Centre for Economic Performance, LSE.
- Head, K., Ries, J., 1998. Immigration and Trade Creation: Econometric Evidence from Canada. *The Canadian Journal of Economics / Revue canadienne d’Economie* 31, 47–62. doi:10.2307/136376
- Helliwell, J.F., 1998. *How Much Do National Borders Matter?* Brookings Institution Press.
- Hoekman, J., Frenken, K., Tijssen, R.J.W., 2010. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy* 39, 662–673. doi:10.1016/j.respol.2010.01.012
- Hunt, J., Gauthier-Loiselle, M., 2008. *How Much Does Immigration Boost Innovation?* (Working Paper No. 14312). National Bureau of Economic Research.
- Jaffe, A.B., 1986. Technological Opportunity and Spillovers of R & D: Evidence from Firms’ Patents, Profits, and Market Value. *The American Economic Review* 76, 984–1001. doi:10.2307/1816464
- Javorcik, B.S., Özden, Ç., Spatareanu, M., Neagu, C., 2011. Migrant networks and foreign direct investment. *Journal of Development Economics* 94, 231–241. doi:10.1016/j.jdeveco.2010.01.012
- Jensen, P.H., Thomson, R., Yong, J., 2011. Estimating the patent premium: Evidence from the Australian Inventor Survey. *Strategic Management Journal* 32, 1128–1138. doi:10.1002/smj.925
- Kapur, D., McHale, J., 2005. *Give us your best and brightest: The global hunt for talent and its impact on the developing world.* Brookings Inst Press.
- Katz, J.S., Martin, B.R., 1997. What is research collaboration? *Research Policy* 26, 1–18. doi:10.1016/S0048-7333(96)00917-1

- Kerr, W.R., 2008. Ethnic Scientific Communities and International Technology Diffusion. *Review of Economics and Statistics* 90, 518–537. doi:10.1162/rest.90.3.518
- Kugler, M., Levintal, O., Rapoport, H., 2013. Migration and Cross-Border Financial Flows (CReAM Discussion Paper Series No. 1317). Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.
- Kugler, M., Rapoport, H., 2005. Skilled Emigration, Business Networks and Foreign Direct Investment (CESifo Working Paper Series No. 1455). CESifo Group Munich.
- Kugler, M., Rapoport, H., 2007. International labor and capital flows: Complements or substitutes? *Economics Letters* 94, 155–162. doi:10.1016/j.econlet.2006.06.023
- Manning, S., Massini, S., Lewin, A.Y., 2008. A Dynamic Perspective on Next-Generation Offshoring: The Global Sourcing of Science and Engineering Talent. *ACAD MANAGE PERSPECT* 22, 35–54. doi:10.5465/AMP.2008.34587994
- Martínez, C., Rama, R., 2012. Home or next door? Patenting by European food and beverage multinationals. *Technology Analysis & Strategic Management* 24, 647–661. doi:10.1080/09537325.2012.705118
- Mayer, T., Zignago, S., 2011. Notes on CEPII's distances measures: the GeoDist database [WWW Document]. URL <http://mpira.ub.uni-muenchen.de/36347/> (accessed 9.2.13).
- McCallum, J., 1995. National Borders Matter: Canada-U.S. Regional Trade Patterns. *American Economic Review* 85, 615–23.
- Migueluez, E., Fink, C., 2013. Measuring the International Mobility of Inventors: A New Database (WIPO Economic Research Working Paper No. 8). World Intellectual Property Organization - Economics and Statistics Division.
- Montobbio, F., 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. doi:10.1016/j.worlddev.2012.11.017
- Ottaviano, G.I.P., Peri, G., Wright, G.C., 2013. Immigration, Offshoring, and American Jobs. *American Economic Review* 103, 1925–1959. doi:10.1257/aer.103.5.1925
- Özden, Ç., Parsons, C.R., Schiff, M., Walmsley, T.L., 2011. Where on Earth is Everybody? *World Bank Economic Review* 25, 12–56.
- Patel, P., Pavitt, K., 1991. Large Firms in the Production of the World's Technology: An Important Case of "Non-Globalisation". *Journal of International Business Studies* 22, 1–21. doi:10.2307/155237
- Patel, P., Vega, M., 1999. Patterns of internationalisation of corporate technology: location vs. home country advantages. *Research Policy* 28, 145–155. doi:10.1016/S0048-7333(98)00117-6
- Picci, L., 2010. The internationalization of inventive activity: A gravity model using patent data. *Research Policy* 39, 1070–1081. doi:10.1016/j.respol.2010.05.007
- Rauch, J., 2003. *Diasporas and Development: Theory, Evidence, and Programmatic Implications*. Department of Economics, University of California at San Diego.
- Rauch, J.E., 2001. Business and Social Networks in International Trade. *Journal of Economic Literature* 39, 1177–1203. doi:10.2307/2698523
- Rauch, J.E., Trindade, V., 2002. Ethnic Chinese Networks in International Trade. *Review of Economics and Statistics* 84, 116–130. doi:10.1162/003465302317331955

- Regets, M., 2001. Research and Policy Issues in High-Skilled International Migration: A Perspective with Data from the United States (IZA Discussion Paper No. 366). Institute for the Study of Labor (IZA).
- Santos Silva, J.M.C., Tenreyro, S., 2006. The Log of Gravity. *Review of Economics and Statistics* 88, 641–658. doi:10.1162/rest.88.4.641
- Santos Silva, J.M.C., Tenreyro, S., 2010. On the existence of the maximum likelihood estimates in Poisson regression. *Economics Letters* 107, 310–312. doi:10.1016/j.econlet.2010.02.020
- Saxenian, A., 1999. Silicon Valley's New Immigrant Entrepreneurs.
- Saxenian, A., 2002. Silicon Valley's New Immigrant High-Growth Entrepreneurs. *Economic Development Quarterly* 16, 20–31. doi:10.1177/0891242402016001003
- Saxenian, A., 2006. *The New Argonauts: Regional Advantage in a Global Economy*. Harvard University Press.
- Scellato, G., Franzoni, C., Stephan, P., 2012. Mobile Scientists and International Networks (Working Paper No. 18613). National Bureau of Economic Research.
- Schmoch, U., 2008. Concept of a technology classification for country comparisons. Final report to the World Intellectual Property Organization (WIPO), Fraunhofer Institute for Systems and Innovation Research, Karlsruhe.
- Thomson, R., de Rassenfosse, G., Webster, E., 2013. Quantifying the effect of R&D offshoring on industrial productivity at home. Mimeo, University of Melbourne.
- Thursby, J.G., Thursby, M.C., 2006. Here or There? A Survey of Factors in Multinational R&D Location: Report to the Government/University/Industry Research Roundtable (SSRN Scholarly Paper No. ID 1456460). Social Science Research Network, Rochester, NY.
- Van Zeebroeck, N., van Pottelsberghe de la Potterie, B., 2011. Filing strategies and patent value. *Economics of Innovation and New Technology* 20, 539–561. doi:10.1080/10438591003668646
- Windmeijer, F.A.G., Silva, J.M.C.S., 1997. Endogeneity in Count Data Models: An Application to Demand for Health Care. *Journal of Applied Econometrics* 12, 281–94.
- World Bank, 2009. World development report. Reshaping Economic Geography. The World Bank, Washington DC.

Appendix 1.

List of developed countries

Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Republic of Korea, Spain, Sweden, Switzerland, United Kingdom, and United States of America.

List of developing/emerging/transition countries

Algeria, Argentina, Armenia, Bangladesh, Belarus, Bosnia and Herzegovina, Brazil, Bulgaria, Cameroon, Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Ecuador, Egypt, Estonia, Ethiopia, Georgia, Ghana, Greece, Guatemala, Hungary, Iceland, India, Indonesia, Iran (Islamic Republic of), Iraq, Israel, Jamaica, Jordan, Kenya, Latvia, Lebanon, Lithuania, Luxembourg, Malaysia, Mauritius, Mexico, Moldova, Morocco, Nepal, Nigeria, Pakistan, Peru, Philippines, Poland, Portugal, Romania, Russian Federation, Singapore, Slovakia, Slovenia, South Africa, Sri Lanka, Syrian Arab Republic, T F Y R of Macedonia, Thailand, Trinidad and Tobago, Tunisia, Turkey, Ukraine, Uruguay, Venezuela, and Viet Nam.

Appendix 2.

Table A.2.1. Summary statistics

	Observations	Mean	St. Dev	Min.	Max.
Collab. inv_i- inv_j	26,380	1.25	10.51	0	678
Collab. app_i- inv_j	26,380	1.71	14.80	0	708
Diaspora size	26,380	22.11	445.32	0	26,661
Diaspora share	26,380	0.00	0.00	0.00	0.05
Distance	26,380	6,875.48	4,539.47	59.62	19,629.50
Contiguity	26,380	0.01	0.10	0	1
Common language	26,380	0.09	0.29	0	1
Lang. similarity	26,380	0.15	0.18	0.00	0.89
Colonial links	26,380	0.04	0.18	0	1
Religion similarity	26,380	0.13	0.19	0.00	0.90
EXP+IMP	26,380	0.01	0.03	0.00	0.41
Tech.distance	26,380	0.48	0.27	0.02	1.00
# patents_i	26,380	810.06	3,487.28	0	64,990
# patents_j	26,380	43,044.81	99,308.57	52	692,364
GDP p.c._i	26,160	9,258.23	9,566.77	432.05	74,113.90
GDP p.c._j	26,380	29,386.74	6,175.00	11,382.60	48,799.70

Notes: 'i' and 'j' stand for, respectively, migrant inventor's sending country and migrant inventor's receiving country'.

Table A.2.2. Correlation matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1															
2	0.93	1														
3	0.37	0.37	1													
4	0.28	0.28	0.82	1												
5	0.01	0.03	-0.12	-0.16	1											
6	0.05	0.02	0.11	0.12	-0.27	1										
7	0.07	0.08	0.13	0.16	0.1	0.05	1									
8	-0.03	-0.03	-0.01	0.01	-0.09	0.07	-0.19	1								
9	0.01	0.02	0.15	0.14	-0.05	0.13	0.32	-0.08	1							
10	-0.03	-0.03	-0.04	-0.03	-0.03	0.12	0.09	0.33	0.15	1						
11	0.11	0.11	0.33	0.2	-0.13	0.08	0.06	-0.02	0.13	0.12	1					
12	-0.15	-0.16	-0.41	-0.41	0.13	-0.07	0.03	-0.07	-0.02	-0.08	-0.24	1				
13	0.21	0.21	0.46	0.47	-0.16	0.1	-0.08	0.07	-0.01	0.03	0.23	-0.73	1			
14	0.18	0.17	0.54	0.21	0.00	0.02	-0.02	-0.03	0.04	-0.07	0.41	-0.21	0.26	1		
15	0.05	0.04	0.14	0.13	-0.23	0.12	-0.07	0.15	0.00	0.17	0.15	-0.44	0.52	0.09	1	
16	0.11	0.11	0.35	0.23	-0.13	0.02	0.00	0.12	-0.03	0.01	0.15	-0.22	0.26	0.51	0.09	1

Notes: 1. Collab. inv_i- inv_j; 2. Collab. app_i- inv_j; 3. ln(Diaspora size); 4. ln(Diaspora share); 5. ln(Distance); 6. Contiguity; 7. Common language; 8. Language similarity; 9. Colonial links; 10. Religion similarity; 11. ln(EXP+IMP); 12. ln(tech. distance); 13. ln(# patents)_i; 14. ln(# patents)_j; 15. ln(GDP p.c.)_i; 16. ln(GDP p.c.)_j. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’.

Figure 1. Share of immigrant inventors over time, 1990-2010, by selected countries

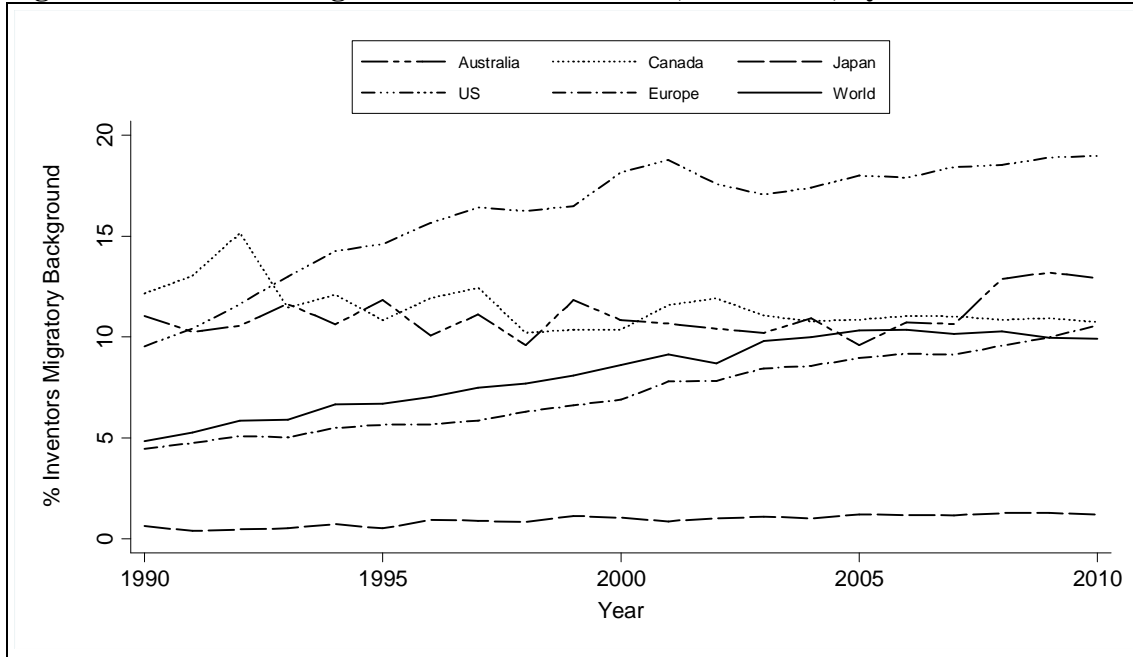


Figure 2. Immigration rates of inventors, 2001-2010, receiving countries

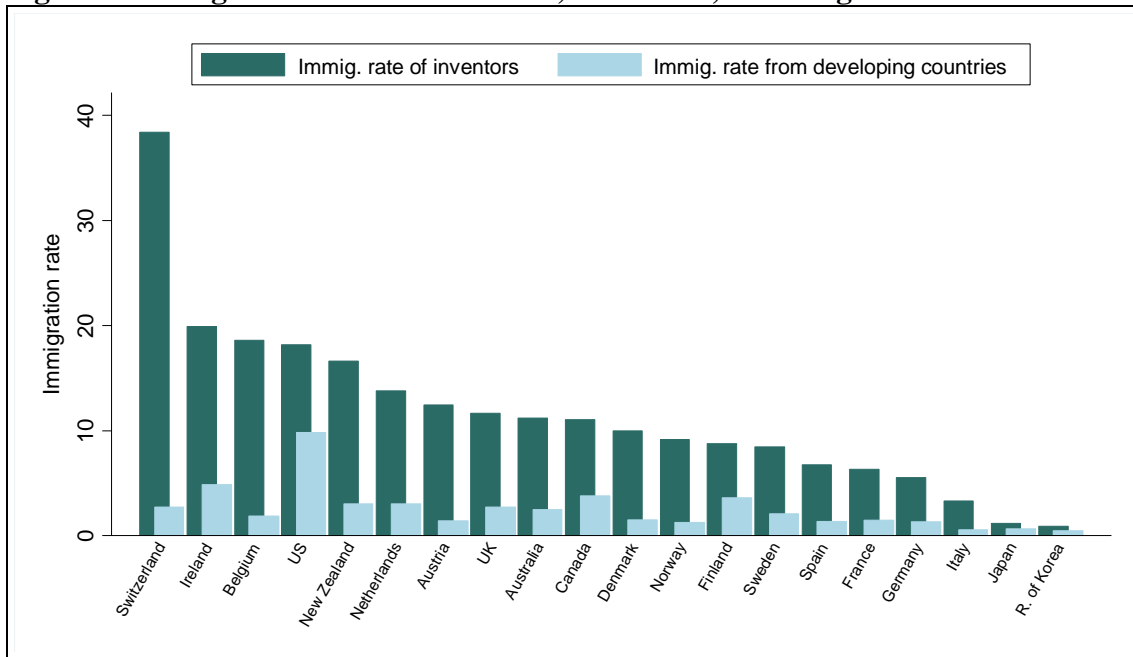


Figure 3. Yearly estimated coefficients of diaspora on co-inventorship, 1991-2010

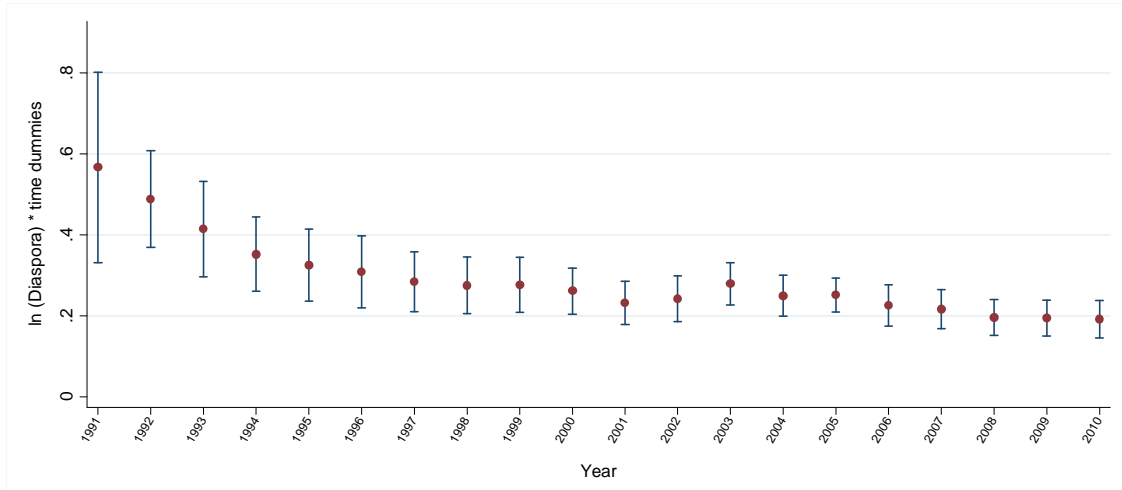


Figure 4. Yearly estimated coefficients of diaspora on R&D offshoring, 1991-2010

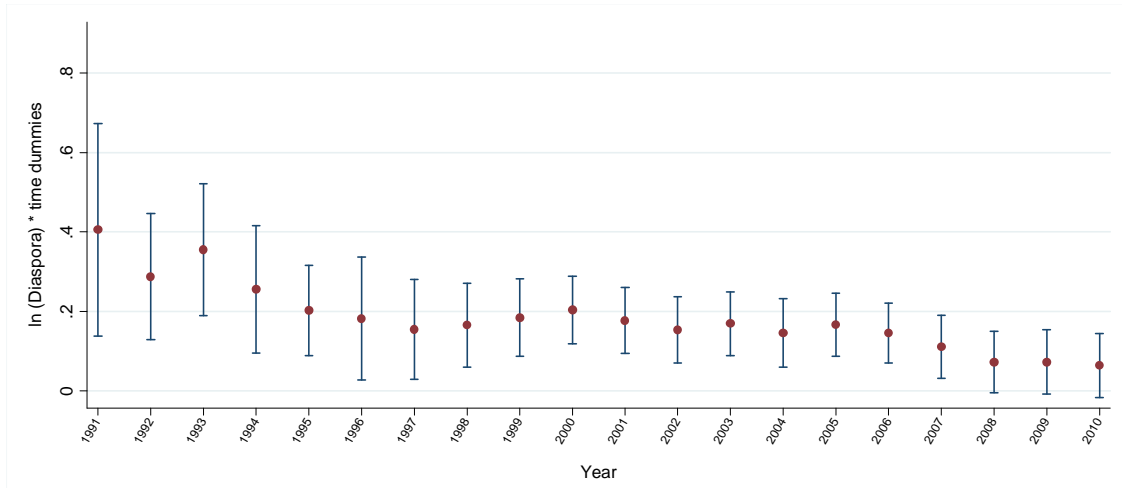


Table 1. Top-20 most populated corridors, 2001-2010

Largest inventor migration corridors			Largest inventor migration corridors, limited to non-OECD sending countries		
Sending	Receiving	Counts	Sending	Receiving	Counts
China	United States	44,444	China	United States	44,444
India	United States	35,607	India	United States	35,607
Canada	United States	18,745	Russia	United States	4,347
U.K.	United States	14,897	China	Japan	2,514
Germany	United States	10,290	China	Singapore	1,925
Germany	Switzerland	8,199	Turkey	United States	1,923
R. of Korea	United States	7,264	Iran	United States	1,442
France	United States	6,540	Romania	United States	1,229
Japan	United States	5,065	Russia	Germany	1,217
Russia	United States	4,347	Mexico	United States	1,164
Australia	United States	3,243	Brazil	United States	1,116
Israel	United States	2,968	Malaysia	Singapore	1,094
France	Switzerland	2,748	Ukraine	United States	977
Netherlands	United States	2,708	China	U.K.	921
Austria	Germany	2,676	China	Germany	889
France	Germany	2,601	India	Singapore	847
China	Japan	2,514	Argentina	United States	821
Italy	United States	2,503	Singapore	United States	771
Germany	Netherlands	2,289	Malaysia	United States	728
Netherlands	Germany	2,140	South Africa	United States	721

Table 2. Baseline specifications: inventor diasporas on co-inventorship and R&D offshoring

	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	PPML	PPML	PPML	PPML	PPML
	Co-inventorship			R&D offshoring		
ln(Diaspora)	0.387*** (0.0268)		0.200*** (0.0229)		0.0929** (0.0407)	
ln(Diaspora share)		0.458*** (0.0353)		0.283*** (0.0255)		0.151*** (0.0468)
ln(Distance)			-0.267*** (0.0523)	-0.282*** (0.0496)	-0.105 (0.0721)	-0.109 (0.0707)
Contiguity			0.210* (0.127)	0.197 (0.126)	-0.0996 (0.211)	-0.0996 (0.214)
Common language			0.625*** (0.132)	0.602*** (0.129)	0.931*** (0.220)	0.910*** (0.219)
Lang. similarity			0.523** (0.225)	0.528** (0.211)	0.822** (0.386)	0.818** (0.378)
Colonial links			0.0654 (0.124)	0.0659 (0.122)	0.299* (0.172)	0.297* (0.177)
Religion similarity			0.706*** (0.256)	0.660*** (0.246)	0.219 (0.446)	0.175 (0.452)
ln(EXP+IMP)			0.0457** (0.0197)	0.0318* (0.0170)	0.0710*** (0.0249)	0.0598** (0.0238)
ln(Tech.distance)			-0.0637 (0.0469)	-0.0632 (0.0460)	-0.245*** (0.0598)	-0.243*** (0.0585)
ln(# patents_i)			0.334*** (0.0530)	0.342*** (0.0462)	0.354*** (0.0693)	0.355*** (0.0660)
ln(# patents_j)			0.0568 (0.115)	0.114 (0.114)	0.333 (0.205)	0.371* (0.209)
ln(GDP p.c. _i)			1.098*** (0.233)	1.112*** (0.188)	1.807*** (0.318)	1.801*** (0.292)
ln(GDP p.c. _j)			-0.0776 (0.552)	-0.583 (0.545)	-1.107 (0.823)	-1.400* (0.812)
Constant	0.546** (0.271)	4.879*** (0.280)	-6.072 (6.006)	0.906 (5.782)	-4.780 (8.651)	-0.901 (8.471)
Observations	26,380	26,380	26,160	26,160	26,160	26,160
Pseudo R2	0.918	0.907	0.958	0.960	0.918	0.918
Sending FE	Yes	Yes	Yes	Yes	Yes	Yes
Receiving FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Log Lik	-20219.27	-20180.31	-18492.28	-18324.09	-23746.92	-23651.86

Notes: Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Per capita GDP at origin presents several missing observations: for 1990, missing data correspond to Azerbaijan, Eritrea, Cambodia and Latvia; for 1991, to Eritrea; and for 2005, to Cyprus, Gabon, Lesotho, Oman, Rwanda, Thailand, Uzbekistan and Zimbabwe. Moreover, data for the former Soviet Republics are only available from 1991; data for TFYR of Macedonia, Croatia and Slovenia only from 1992; and data for the Czech Republic, Slovakia and Eritrea only from 1993. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’.

Table 3. Baseline specifications with time-varying multilateral resistance

	(1)	(2)	(3)	(4)
	PPML	PPML	PPML	PPML
	Co-inventorship		R&D offshoring	
ln(Diaspora)	0.243*** (0.0233)		0.111*** (0.0423)	
ln(Diaspora share)		0.260*** (0.0246)		0.104** (0.0440)
ln(Distance)	-0.0468 (0.0552)	-0.0959* (0.0558)	0.173** (0.0760)	0.150** (0.0764)
Contiguity	0.116 (0.122)	0.0872 (0.121)	-0.271 (0.191)	-0.287 (0.191)
Common language	0.434*** (0.108)	0.467*** (0.106)	0.703*** (0.187)	0.725*** (0.185)
Lang. similarity	0.405** (0.202)	0.445** (0.189)	0.694** (0.352)	0.720** (0.346)
Colonial links	0.0315 (0.103)	0.0568 (0.106)	0.291** (0.143)	0.303** (0.141)
Religion similarity	0.612*** (0.235)	0.617*** (0.231)	0.185 (0.419)	0.211 (0.415)
ln(EXP+IMP)	0.228*** (0.0335)	0.219*** (0.0347)	0.329*** (0.0495)	0.331*** (0.0507)
ln(Tech.distance)	-0.136*** (0.0506)	-0.150*** (0.0505)	-0.287*** (0.0689)	-0.292*** (0.0690)
Constant	-0.410 (0.652)	2.399*** (0.686)	-1.726** (0.791)	-0.566 (0.906)
Observations	19,676	19,676	20,574	20,574
Pseudo R2	0.978	0.978	0.956	0.956
Sending FE	Yes	Yes	Yes	Yes
Receiving FE	Yes	Yes	Yes	Yes
Year FE	No	No	No	No
Sending FE*Time FE	Yes	Yes	Yes	Yes
Receiving FE*Time FE	Yes	Yes	Yes	Yes
Log Lik	-16226.61	-16213.81	-20463.96	-20484.05

Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’. The different number of final observations between Tables 2 and 3 is due to the inclusion of fixed effects in pseudo-maximum likelihood estimations: the PPML method automatically drops the country-specific fixed-effects (and their corresponding observations) for which the country has zero recorded inventors’ flows to every other country in the sample in order to achieve convergence. Results are comparable to other count data methods without removing these observations (see Santos Silva and Tenreyro, 2010, for further details).

Table 4. Inventor diaspora effect between culturally closer/more distant countries

	(1)	(2)	(3)	(4)	(5)
	PPML	PPML	PPML	PPML	PPML
	Co-inventorship				
ln(Diaspora)	0.204*** (0.0247)	0.199*** (0.0229)	0.198*** (0.0230)	0.201*** (0.0230)	0.206*** (0.0249)
ln(Distance)	-0.263*** (0.0535)	-0.270*** (0.0517)	-0.267*** (0.0520)	-0.275*** (0.0560)	-0.262*** (0.0557)
Contiguity	0.213* (0.127)	0.217* (0.128)	0.233* (0.127)	0.216* (0.127)	0.250** (0.127)
Common language	0.690*** (0.193)	0.613*** (0.132)	0.630*** (0.131)	0.628*** (0.131)	0.776*** (0.198)
Lang. similarity	0.508** (0.227)	0.781*** (0.273)	0.528** (0.224)	0.532** (0.224)	0.858*** (0.285)
Colonial links	0.0502 (0.131)	0.0774 (0.124)	0.473* (0.248)	0.0724 (0.126)	0.468* (0.253)
Religion similarity	0.702*** (0.254)	0.676*** (0.262)	0.692*** (0.256)	0.782*** (0.267)	0.639** (0.287)
ln(Dia.)*Com. language	-0.0122 (0.0223)				-0.0302 (0.0241)
ln(Dia.)*Language sim.		-0.0829 (0.0522)			-0.117* (0.0606)
ln(Diaspora)*Colonial			-0.106** (0.0498)		-0.110** (0.0494)
ln(Diaspora)*Religion				-0.0526 (0.0708)	0.00141 (0.0758)
ln(EXP+IMP)	0.0456** (0.0196)	0.0444** (0.0195)	0.0457** (0.0197)	0.0445** (0.0199)	0.0438** (0.0197)
ln(Tech.distance)	-0.0662 (0.0468)	-0.0663 (0.0469)	-0.0647 (0.0465)	-0.0624 (0.0475)	-0.0748 (0.0464)
ln(# patents_i)	0.338*** (0.0532)	0.327*** (0.0525)	0.335*** (0.0535)	0.332*** (0.0539)	0.335*** (0.0517)
ln(# patents_j)	0.0540 (0.114)	0.0614 (0.116)	0.0356 (0.116)	0.0632 (0.116)	0.0339 (0.115)
ln(GDP p.c. _i)	1.085*** (0.231)	1.101*** (0.234)	1.087*** (0.235)	1.084*** (0.239)	1.062*** (0.237)
ln(GDP p.c. _j)	-0.0716 (0.552)	-0.100 (0.550)	0.0349 (0.551)	-0.103 (0.554)	0.0227 (0.550)
Constant	-6.102 (6.019)	-5.847 (6.000)	-6.941 (5.985)	-5.701 (6.080)	-6.745 (6.044)
Observations	26,160	26,160	26,160	26,160	26,160
Pseudo R2	0.958	0.958	0.957	0.957	0.958
Sending FE	Yes	Yes	Yes	Yes	Yes
Receiving FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Log Lik	-18490.83	-18482.05	-18478.32	-18490.28	-18459.73

Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’.

Table 5. GMM estimates with instrumented diaspora

	(1)	(3)
	GMM	GMM
	Co-inventorship	R&D offshoring
ln(Diaspora)	0.223*** (0.0774)	0.120 (0.170)
ln(Distance)	-0.247*** (0.0649)	-0.132 (0.0985)
Contiguity	0.539** (0.220)	1.058*** (0.388)
Common language	0.630** (0.261)	0.199 (0.543)
Lang. similarity	0.215* (0.128)	-0.0970 (0.210)
Colonial links	0.622*** (0.133)	0.973*** (0.251)
Religion similarity	0.0612 (0.122)	0.269 (0.180)
ln(EXP+IMP)	0.0535** (0.0260)	0.0620 (0.0388)
ln(Tech.distance)	-0.0779* (0.0451)	-0.236*** (0.0618)
Constant	-6.866 (5.998)	-0.0369 (8.833)
Observations	26,160	26,160
Controls	Yes	Yes
Sending FE	Yes	Yes
Receiving FE	Yes	Yes
Year FE	Yes	Yes
F-test	344.58	344.58
p-value	0.0000	0.0000
Hansen's J chi2	3.00468	6.13152
p-value	0.3909	0.1054

Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’. Instruments are centered around their mean, and they are as follows: (i) the size of the bilateral diaspora between countries i and j in the 1960s – and its square, and (ii) the size of the unskilled diaspora original from country i residing in country j in 1990 (migrants with only primary education) – and its square.

Table 6. The effect of inventor diasporas, by technology field

	(1)	(2)	(3)	(4)	(5)
	PPML	PPML	PPML	PPML	PPML
	Electrical engineering	Instruments	Chemistry	Mechanical	Other sectors
Co-inventorship					
ln(Diaspora)	0.248*** (0.0437)	0.180*** (0.0294)	0.161*** (0.0248)	0.149*** (0.0364)	0.183*** (0.0420)
ln(Distance)	-0.139 (0.0850)	-0.402*** (0.0672)	-0.319*** (0.0673)	-0.385*** (0.0774)	-0.696*** (0.122)
Constant	-31.16*** (11.71)	-8.392 (11.76)	7.643 (5.602)	11.72 (9.777)	33.93* (18.85)
Observations	23,500	24,300	25,760	23,780	20,740
Pseudo R2	0.932	0.889	0.912	0.796	0.699
Sending FE	Yes	Yes	Yes	Yes	Yes
Receiving FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R&D offshoring					
ln(Diaspora)	0.203*** (0.0665)	0.129*** (0.0331)	0.0892** (0.0402)	0.152*** (0.0415)	0.106 (0.0709)
ln(Distance)	0.0618 (0.123)	-0.216** (0.0934)	-0.0473 (0.105)	-0.175 (0.123)	-0.285 (0.173)
Constant	-8.768 (17.39)	14.61 (10.47)	0.175 (8.569)	9.154 (12.28)	17.99 (21.34)
Observations	24,700	24,960	25,760	24,560	21,140
Pseudo R2	0.872	0.884	0.838	0.766	0.505
Sending FE	Yes	Yes	Yes	Yes	Yes
Receiving FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’.

Table 7. Are China and India the exception rather than the rule?

	(1)	(2)	(3)	(4)
	PPML	PPML	PPML	PPML
	Co-inventorship			
	No BRICS	No US	No BRICS, no US	No BRICS, no US
ln(Diaspora)	0.192*** (0.0384)	0.204*** (0.0358)	0.200*** (0.0419)	0.190*** (0.0465)
ln(Distance)	-0.447*** (0.0631)	-0.325*** (0.0582)	-0.515*** (0.0654)	-0.269*** (0.0759)
Contiguity	0.0729 (0.138)	0.139 (0.143)	0.102 (0.161)	0.108 (0.155)
Common language	0.703*** (0.146)	0.434*** (0.137)	0.453*** (0.149)	0.425*** (0.142)
Lang. similarity	0.610* (0.328)	0.220 (0.213)	0.367 (0.288)	0.282 (0.283)
Colonial links	0.0905 (0.157)	0.162 (0.114)	0.235* (0.138)	0.112 (0.134)
Religion similarity	0.433* (0.223)	0.779*** (0.251)	0.463** (0.226)	0.443** (0.210)
ln(EXP+IMP)	0.0432* (0.0259)	0.00145 (0.0171)	-0.00173 (0.0211)	0.217*** (0.0411)
ln(Tech.distance)	-0.123** (0.0547)	-0.00597 (0.0668)	0.0560 (0.0697)	0.0595 (0.0824)
ln(# patents_i)	0.130** (0.0538)	0.319*** (0.0575)	0.134** (0.0599)	
ln(# patents_j)	0.128 (0.110)	0.00645 (0.113)	0.0675 (0.111)	
ln(GDP p.c. _i)	0.604* (0.363)	1.038*** (0.233)	0.389 (0.247)	
ln(GDP p.c. _j)	0.115 (0.680)	-0.0884 (0.587)	0.158 (0.697)	
Constant	-5.626 (7.610)	-7.494 (5.929)	-5.900 (7.425)	1.753*** (0.667)
Observations	24,180	24,852	22,971	14,752
Pseudo R2	0.900	0.828	0.658	0.725
Sending FE	Yes	Yes	Yes	No
Receiving FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	No
Sending FE*Time FE	No	No	No	Yes
Receiving FE*Time FE	No	No	No	Yes
Log Lik	-14236.62	-15541.05	-11836.79	-10463.32

Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’.

Table 8. Robustness checks. Co-inventorship

	(1)	(2)	(3)	(4)
	OLS [ln(co-inv.+1)]	NegBin	Zero-inflated Poisson	Zero-inflated NegBin
	Co-inventorship			
ln(Diaspora)	0.232*** (0.0136)	0.221*** (0.0267)	0.261*** (0.0254)	0.256*** (0.0284)
ln(Distance)	-0.0136 (0.0156)	-0.453*** (0.0486)	-0.109* (0.0566)	-0.280*** (0.0607)
Contiguity	0.493*** (0.159)	0.0845 (0.186)	0.248** (0.115)	0.0482 (0.157)
Common language	0.00615 (0.0447)	0.608*** (0.138)	0.619*** (0.136)	0.686*** (0.147)
Lang. similarity	0.00779 (0.0611)	0.536*** (0.203)	0.504** (0.230)	0.607*** (0.203)
Colonial links	-0.0225 (0.0510)	0.455*** (0.140)	-0.141 (0.116)	0.159 (0.145)
Religion similarity	0.0472 (0.0421)	0.370* (0.189)	0.560* (0.311)	0.0421 (0.284)
ln(EXP+IMP)	-0.00681*** (0.00224)	0.0191 (0.0137)	0.0479** (0.0193)	0.0198 (0.0155)
ln(Tech.distance)	-0.0174 (0.0244)	-0.0175 (0.0502)	-0.0308 (0.0417)	-0.0201 (0.0529)
ln(# patents_i)	0.0863*** (0.00991)	0.257*** (0.0367)	0.340*** (0.0534)	0.295*** (0.0436)
ln(# patents_j)	-0.00406 (0.0134)	0.105 (0.0902)	-0.227* (0.123)	-0.195* (0.107)
ln(GDP p.c. _i)	0.249*** (0.0491)	0.957*** (0.168)	1.050*** (0.219)	0.845*** (0.219)
ln(GDP p.c. _j)	-0.232*** (0.0746)	0.209 (0.523)	-0.603 (0.635)	-0.408 (0.687)
Constant	1.089 (0.835)	-6.659 (5.554)	1.449 (6.603)	2.203 (6.881)
Observations	26,160	26,160	26,160	26,160
Sending FE	Yes	Yes	Yes	Yes
Receiving FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Country-pair clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ‘_i’ and ‘_j’ stand for, respectively, migrant inventor’s sending country and migrant inventor’s receiving country’.