Groupe de Recherche en
Économie Théorique et Appliquée

# Highly skilled migration and the internationalization of knowledge

***Claudia NOUMEDEM TEMGOUA***

*GREThA, CNRS, UMR 5113*
*Université de Bordeaux*

*Claudia.noumedem-temgoua@u-bordeaux.fr*

***Cahiers du GREThA***
**n° 2018-16**
**July**

**GRETHA UMR CNRS 5113**
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

## Migrations de personnes hautement qualifiées et internationalisation de connaissances

### Résumé

*Ce papier aborde la question du rôle des diasporas constituées de personnes hautement qualifiées d'origine chinoise et indienne dans l'internationalisation des réseaux de connaissances, pour un échantillon de pays de destination membres de l'OCDE. Plus précisément, deux principaux types de réseaux de connaissances sont analysés: les réseaux de co-inventeurs et de co-auteurs. Des données à l'échelle nationale sur les migrants hautement qualifiés tirées de la base de données OCDE-DIOC (Base de données sur les immigrés dans les pays de l'OCDE, 2000/01, 2010/11) et des informations sur les réseaux de co-auteurs et de co-inventeurs provenant des publications et des brevets sont utilisées conjointement. L'analyse de ce chapitre est basée sur des régressions de modèle de gravité. Les résultats obtenus à l'issu de ces régressions montrent que les pays de destination avec une part importante des diasporas indiennes ou chinoises de personnes hautement qualifiées ont tendance à collaborer davantage sur les publications et les brevets. En étendant l'analyse à d'autres pays, des résultats similaires sont obtenus pour le cas du Vietnam, du Pakistan et de l'Iran.*

**Mots-clés:** Migrations, personnes hautement qualifiées, publications, Coopération en R&D, diffusion, brevets

## Highly skilled migration and the internationalization of knowledge

### Abstract

*This paper investigates the role of Chinese and Indian highly skilled diaspora in the internationalization of knowledge networks, for a sample of OECD destination countries. We mainly focus on two types of knowledge networks: co-inventorship and co-authorship. We jointly exploit country-level data on highly skilled migration and information on co-authorship and co-inventorship from publication and patent data. Based on a gravity model regression analysis, we find that OECD country pairs hosting sizeable portions of the Indian or Chinese highly skilled diasporas tend to collaborate more on publications and patents, after controlling for other migration trends. When extending the analysis to other countries, we find similar results for Vietnam, Pakistan and Iran.*

**Keywords:** migration, highly skilled, publications, R&D cooperation, diffusion, patent

**JEL:** C8, F22, J61, O31, O33

**Reference to this paper: Claudia NOUMEDEM TEMGOUA** (2018) Highly skilled migration and the internationalization of knowledge, *Cahiers du GREThA*, n°2018-16.

http://ideas.repec.org/p/grt/wpegrt/2018-16.html.

**GRETHA UMR CNRS 5113**
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

## 1 Introduction*

A growing literature has been dealing with the role of highly skilled – *hs* – international migrants as a channel of knowledge exchange and circulation across countries and regions. This literature has exploited the underlying idea that sharing a common social and cultural background could favour different types of exchanges within specific diaspora groups. In particular, a common social and cultural background could support the formation and maintenance of social networks of *hs* migrants, where knowledge would be exchanged or circulate more easily, both within the migrants' destination countries and to their countries of origin. This assumption rests on the principle that scientific and technical knowledge contains tacit elements, whose transfer demand direct human interaction and some form of proximity (geographical, cultural…) (Breschi & Lissoni, 2009; Dosi, 1988; Jaffe & Caballero, 1993; Leonard & Sensiper, 1998). Some studies have shown that *hs* migrants from specific origin countries – most notably Indian and Chinese – form a strongly conntected diaspora insofar that they tend to have a higher propensity to pass on knowledge to other *hs* migrants of same origin at the destination than with nationals (Agrawal et al., 2006; Breschi et al., 2015; Kerr, 2009). These findings highlight the importance of interactions or links among *hs* diaspora members of same origin. The set of these interactions form what we call *hs* diaspora knowledge networks. It is within such networks that part of new ideas and innovations are created within destination countries, thus contributing to their economic growth (Ackers, 2005; Agrawal et al., 2011; Gill, 2005; Kerr, 2008).

We observe, however, that there is no reason to presume that social interactions between same-origin migrants ought to be bound to the countries of destination or within the origin-destination axis. In fact, *hs* diaspora members might have a higher propensity to collaborate wherever they are, including across multiple destination countries. In other words, *hs* diaspora knowledge networks may span far beyond one single destination country to two or more destinations. There have been some qualitative studies on the functioning of *hs* diaspora knowledge networks – particularly from developing countries – across destination countries depicting how they come together in an associative platform in order to channel knowledge back home (Adepoju et al., 2008; Brown, 2002; J. P. Meyer & Wattiaux, 2006). But, to the best of our knowledge, no empirical study has yet investigated how *hs* diasporas contribute to the internationalization of knowledge across destination

* **Acknowledgements :** This study has received financial support from the French State in the frame of the "Initiative of Excellence" Programme IdEx Bordeaux.

countries. Moreover, little research has been done on the mechanisms underpinning knowledge transfer among destination countries within immigrant networks.

We thus intend to fill the existing gap in the literature by assessing the impact on collaboration in innovative/knowledge activities of two large *hs* diasporas – Chinese and Indian. We refer to Science & Technology (S&T) collaboration for a large sample of the Organisation for Economic Co-operation and Development (OECD) destination countries. In particular, we explore two variables as proxies for S&T collaboration: co-inventorship and co-authorship, as computed from various data sources. For each of these dependent variables, we apply a gravity approach at a country pair level. We run PPML – Poisson Pseudo Maximum Likelihood – regressions (Santos-Silva & Tenreyro, 2006) for each of the two variables. Our preliminary results suggest a positive impact of these *hs* diasporas on all our knowledge collaboration variables. Additionally, we test for the impact of other *hs* diasporas, all from top *hs* migrant sending countries to OECD host countries, and find positive results for Vietnam, Pakistan and Iran. Although the number of Indian and Chinese *hs* diasporas are a way larger, these other *hs* diasporas produce similar effects on co-inventorship and co-authorship.

The rest of the paper is organized as followed: in section 2 we present a review of the literature, in section 3 we briefly discuss key definitions, while in section 4 we develop our methodology. In section 5 we discuss our results, and finally in the last section we conclude.

## *2 Literature review*

### *2.1 Migration, social networks and innovation*

Traditional studies on international migration have been conducted either as part of development economics or within the framework of labour economics. The origins of this approach can be traced back to basic neoclassical models establishing a potential for considerable efficiency gains from a more liberal international mobility of labour, (Klein & Ventura, 2007; Moses & Letnes, 2004). Further theories have linked migrants to human capital formation and wages in receiving countries (Massey et al., 1993); but also to financial remittances, education and growth in sending countries (M. Beine et al., 2001; Mountford, 1997; Stark & Wang, 2002).

In parallel, the original neoclassical models also stand as the basic framework to the growing body of theoretical and empirical literature that has explored the role of migrants in favouring transactions between countries. This literature has emphasized the externalities derived from migrant networks in terms of the social and economic linkages between their home and destination countries. These migrant networks externalities act indirectly in reducing informal barriers; and so lowering transaction costs in bilateral economic exchanges between countries. A strand of studies has documented a positive impact of migrant networks on bilateral FDI (Javorcik et al., 2011; Kugler & Rapoport, 2007), firms' internationalization strategies (Foley & Kerr, 2013; Saxenian et al., 2002), international knowledge diffusion (Agrawal et al., 2011; Kerr, 2008) and trade (Gould, 1994).

There has been extensive work on the role played by migrants in boosting bilateral trade (Dunlevy, 2006; Felbermayr et al., 2010; Herander & Saavedra, 2005). Most of these studies use gravity models to assess the pro-trade impact of direct migrant connections between home and host countries along two channels: the preference and the trade-cost ones. The first channel is related to the level of utility migrants might derive from certain goods as compared with others. Thus they will tend to trade more goods from which they get a higher utility in their host countries (Girma & Yu, 2000; Gould, 1994; Head & Ries, 1998; Wagner & Leydesdorff, 2005). The second channel is a self-enforcing mechanism through which migrant networks may help overcoming informal barriers – for instance language, culture and institutions and favour the creation and strengthening of business relationships. Migrants may also carry with them valuable information on foreign business opportunities (Dunlevy, 2006; Herander & Saavedra, 2005). This second channel appears to be the most relevant for explaining other types of international exchanges such as FDI and knowledge, on which migrant connections have been found to have a direct impact.

An interesting development of the literature on migration and trade has explored the role of what we will refer to as *indirect* migrant connections, which connect minorities from the same origin country across different destinations (Felbermayr et al., 2010; Giovannetti & Lanati, 2015). The seminal work by (Rauch & Trindade, 2002) is considered as the first empirical work to explore this question, with a special attention on Chinese migrants. The results from this study predict a large indirect trade creation effect of Chinese migrants in their host countries. More precisely, the authors find a large and strong effect: the presence of Chinese population share in two countries, at the levels that prevail in South East Asia leads to an estimated average increase of at least 60 percentage points in bilateral trade in differentiated products between these countries. In general, similarly to the direct migrant connections impacts, the pro-trade effects of indirect migrant networks are not just determined by preferences for certain goods, but also by an alleviation of information frictions. By extension, the literature has investigated other types of international transactions or collaborations enabled by migrant networks, such as knowledge exchanges related to innovation activities.

Linking migration to innovation or international knowledge diffusion has long been considered challenging until the recent development of new global-scale micro data from a variety of sources. This has resulted in an increasing amount of empirical production addressing various issues on the role of migration in innovation or knowledge diffusion in both sending and receiving countries. These studies mainly focus on the specific category of *hs* migrants, in particular those with degrees or jobs in Science, Technology, Engineering or Mathematics (STEM) (Breschi & Lissoni, 2009; Chellaraj et al., 2008; Hunt & Gauthier-Loiselle, 2010; Kerr, 2009). The most common data sources include labour force surveys and censuses at a national as well as at a global level (Docquier et al., 2007; Docquier & Rapoport, 2012). More recently, a new body of literature has emerged, which uses bibliometric data to track the international mobility of researchers (Appelt et al., 2015; Conchi & Michels, 2014; Kamalski & Plume, 2013; Laudel, 2003; Moed & Halevi, 2014; Moed & Plume, 2013; Pierson & Cotgreave, 2000). Finally, patent data have also been exploited, due to three attractive features:

- first they provide information on homogenous group of *hs* workers, namely the inventors reported on patent applications;
- second they make it possible to identify migrants by comparing information on the inventors' residence, as reported on the patent documentation, and either their nationality, which may be also reported on the documentation or inferred with the help of name analysis techniques;
- third, they can help to capture international innovation or knowledge diffusion by means, respectively, of cross-country co-patenting and patent citation analysis.

A study by (Miguelez, 2016) stands out as a good illustration of this triple advantage of patent data. The author uses inventor data from PCT (Patent Cooperation Treaty) patent records issued by the World International Patent Organization (WIPO). He investigates the impact of *hs* diasporas on the globalisation of R&D activities. Information on the migrant status of inventors is obtained by comparing the inventors' nationality to their residence at the time of patent filing. As for cross-country collaboration in technology, patent data allow for two measures: co-inventorship and R&D offshoring. The author applies a gravity model with country pairs as observations and either one of the two measures as the dependent variable. The focal regressor is the stock of active *hs* diaspora from one country *j* in a host country *j*, which turns out to have a strong and positive impact on both dependent variables. More precisely, the author finds a 10% increase in the inventor diaspora from *i* in *j* leads to an increase of around 2% in international patent collaborations at the level of inventors. This paper represents one major contribution to the empirical literature on the role of direct migrant connections on knowledge flows. However, the focal point in Miguelez' paper remains knowledge diffusion from destination to origin countries and not across destination countries.

In the migration and innovation literature in general, most empirical studies have focused on questions related to the role played by *hs* diasporas in the diffusion of knowledge to their origin countries (Kerr, 2008) or in enhancing innovation within each specific destination country (Breschi et al., 2014; Chellaraj et al., 2008; Ottaviano & Peri, 2006). This means there are still some unaddressed questions such as how same origin *hs* diasporas contribute to knowledge diffusion or exchange across destination countries. Besides, this body of literature has been strongly dominated by studies on the US as a host country, although with few exceptions (Niebuhr, 2010; Ozgen et al., 2013).

### *2.2 Science and technology (S&T) collaborations*

Bibliometric and patent data have been some of the most widely used data to help investigating collaboration patterns both across individuals – e.g. Scientists, authors and inventors –, institutions, and at a more global scale across countries or regions (Abramo et al., 2012; Kamalski & Plume, 2013; Luukkonen et al., 1993; Narin et al., 1991; Wagner, 2005; Wagner & Leydesdorff, 2005; Yoshikane & Kageura, 2004). Indeed, information recorded in publications on authors and their affiliations have made it possible to develop some measures of scientific collaboration, based upon co-authorship. Besides, publications are one of the most common means of documented scientific communication and collaboration. As such, they have the advantage of covering a wider range of sectors. They therefore appear as an appropriate data source for studies on international collaboration across sectors – university, firm and government – (Glänzel & Schubert, 2004). As for patent data, they are more connected to the industrial R&D. That partly explains why they have been widely used in studies on the link between firms' alliances and innovation, along with data on R&D cooperation (Ponds et al., 2007). This implies each of the above types of data has its own embedded specificities, which may help capturing certain patterns of collaboration better than others. Yet, a common feature to all collaboration patterns is that they are informative of the existence of knowledge networks or linkages that may expand beyond sectors, institutional and national borders, with potential benefits to the involved parties.

Overall, there has been an increasing tendency to international collaborations over the past years worldwide and in OECD countries in particular (Guellec & de la Potterie, 2001). This is due to factors such as massive funding, an increasing mobility of researchers and changes in communication patterns particularly in the scientific world (Glänzel & Schubert, 2004). Also, over the last decades global firms have moved their main technology creation activities from a home-country based perspective to a more internationally-oriented one. This is partly explained by the surge of innovation and rapid trends and requirements of market trends, thus the need to exploit knowledge and technology from different sources abroad (Bastian, 2006; Bresson, 1996; De Backer & Basri, 2008; Nooteboom, 1999; Von Hippel, 1988). Besides these market oriented motivations, there are additional political and institutional factors – such as the European initiative for European collaboration and the Framework Programme for research and technological development at the European level – aiming at promoting knowledge creation and more particularly R&D cooperation at a regional level (Removille & Clarysse, 1999).

In general firms are the leading players in R&D networks formation or collaboration agreements (De Backer & Basri, 2008). Consequently, studies on the determinants of the globalization of knowledge and technology have long been confined to the firm level (Granstrand et al., 1993). Moreover, incomplete harmonized data on countries' cross-border R&D flows has rendered it difficult to perform cross-countries analysis. However, in recent years new strategies have been used to counter this data shortage. One of these strategies has been to use co-patenting or co-inventorship recorded in patent data at a country level as a proxy for cross-country R&D collaboration. Admittedly, it is important to point out to the singularity of the latter strategy insofar that co-patenting or co-inventorship only representing outcomes from R&D activities, the entire picture might not be fully depicted. Furthermore, unlike other types of S&T collaborations, those captured by patent data mainly involve research collaboration for market purposes and industry-oriented activities (Maggioni et al., 2007). As a result, applied knowledge is often the main outcome of such collaborations (Lata et al., 2012).

Yet, patent data have been increasingly used in various studies on the determinants of international collaboration in innovation. For instance, (Guellec & de la Potterie, 2001) study the determinants of internationalisation of knowledge at the country level, as proxied by three indicators, namely the shares of patents with the joint presence of: a domestic and a foreign inventor; a foreign inventor and domestic applicant; and a domestic inventor-foreign applicant pairs. They find the internationalisation of a country's technological activity decreases with that country GDP and R&D intensity. This can be interpreted as follows: researchers from large countries prefer to tap into their country's own resources for technology generation. In other words these researchers find it easier to collaborate with other researchers who are closer to them as their high level of technology does not raise any incentive for them to look for any knowledge source elsewhere. The UK and the US are the exception to this rule and (Guellec & de la Potterie, 2001) explain it with their language similarity with many countries, which thus makes researchers from these countries likely to cooperate more easily with other countries. The latter point implies that language similarity favours the internationalisation of knowledge, as suggested by their results. They also find geographical and technological proximity to foster bilateral cooperation in technology.

As for scientific collaboration, it has been mainly carried out by research institutions such as universities, public and private research centres. The principal motivations of its main actors are among others access to funding and equipment, access to expertise, speeding up progress, enhancing productivity, and reducing isolation (Beaver, 2001). Its study can be traced back to the 1960s (Clarke, 1967). However, most of the research on this topic has used a descriptive approach, at least until recent years (Glänzel et al., 1999; Jones et al., 2008; Katz, 1994; Narin et al., 1991; Okubo & Zitt, 2004; Wagner & Leydesdorff, 2005). Indeed, more recently there have been a growing number of empirical studies on the determinants of international research collaboration using bibliometric data. For instance, (Hoekman et al., 2010) who analyze co-publication patterns among European regions for the period 2000-2007. They find patterns of cross-border collaborations to be geographically localized. That is, physical distance impedes on co-publication activities while language similarity favours them.

In general, S&T collaborations – mainly joint publications and joint patents as the ones that are easily traceable with more comprehensive data – have been positively associated with knowledge production and dissemination. For instance, there is a great deal of empirical evidences on the positive impact of R&D alliances on innovation performances (Belderbos et al., 2004; Cincera et al., 2003; Faems et al., 2005). However, most of this literature is at a micro or firm level of analysis and fits more into the scope of management studies, with a special attention on company or university strategies.

All in all, the above theoretical and empirical literature review reveals although *hs* migration has been the topic of a large empirical and theoretical amount of work, linking it to international collaboration in knowledge is a new approach with several questions which are yet to be explored. We therefore intend to contribute to this field of literature by digging into some of these questions.

## 3 Key definitions

Before going further into analysis, we need to introduce the key definitions we will use in the present study. We make a distinction between general definitions – those derived from the migration literature – and specific definitions – which we have elaborated ourselves for our own research purposes.

- ***General definitions***

In general, by "*migrants*" we refer to individuals residing outside their countries of origin at a given point of time. In turn, the migration literature alternatively defines the country of origin according to three non-mutually exclusive criteria: nationality, place of birth and duration of stay. There are specific limitations and advantage tied to each criterion, and often time the choice between them depends on data availability. Our data – a detailed description of which is provided below – come from national labour force surveys and censuses, which mainly refer to migrants' place of birth, with the exception of a couple of sources which use the nationality instead. In general, they do not distinguish between temporary and permanent migration in our analysis[1].

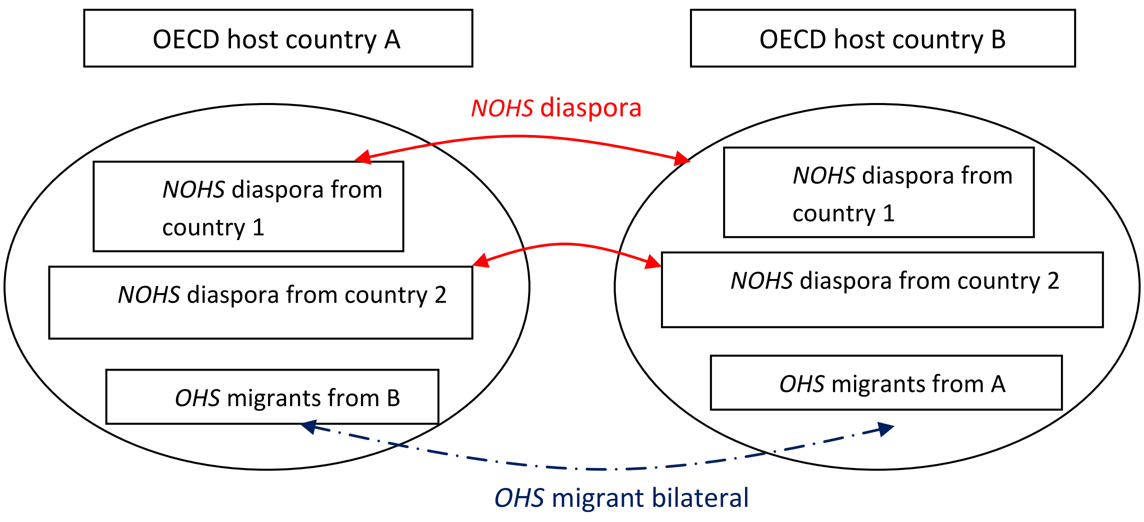
Secondly, drawing from the conceptual definition of the United Nations Educational, Scientific and Cultural Organization (UNESCO) and the OECD Frascati Manual, we identify highly-skilled migrants on the basis of their educational attainment level. More precisely, we define as *hs* individuals having completed a tertiary education level.

[1]Notice that the United Nations (UN) has recommended considering as migrants only those with at least a one-year stay abroad. We acknowledge this recommendation, but our data does not allow is to take into account any duration of stay. However, we do believe not accounting for the time dimension of mobility wouldn't bias our results, as we assume the duration of stay has a minor role to play in the formation of networks

- ***Specific definitions***

Given the research objectives of this paper, we adopt the following key defintions, as illustrated by Figure 1. By non-OECD *hs* (*NOHS*) diaspora – migrants from countries 1 and 2 in our illustration –, we mean all *hs* migrants whose country of origin does not belong to the OECD host countries – represented by country A and B in Figure 1. We are particularly interested about *NOHS* groups of people from the same country of origin, but dispersed across two or more OECD host countries. Migrants in such groups might interact with each other across borders and form connections – as represented by the red bold arrows in Figure 1. These connections differ from the OECD *hs* (*OHS*) migrant bilateral links formed by groups of *hs* migrants originating from and moving to OECD host countries, as depicted by the blue broken arrows.

**Figure 1**: *NOHS* diasporas Vs. *OHS* migrant bilateral links.




Besides, we define S&T collaborations the joint efforts of groups of organisations in performing R&D activities (Lata et al., 2012). Here we explore two types of S&T collaborations at an international dimension; co-inventorship and co-authorship.

# 4 Methodology

## 4.1 Empirical approach

For assessing the contribution of *NOHS* diasporas to international collaboration in S&T, we use a gravity model, which explains the intensity of interaction between two elements with both their individual characteristics and mutual distance. In social sciences, and economics in particular, it has been widely applied at the regional or country level to explain both trade volumes (Bergstrand, 1985; Felbermayr & Toubal, 2012; Rauch, 1999; Tinbergen, 1962), migration (Beine, et al., 2016; Mayda, 2010; Ortega & Peri, 2013) and knowledge diffusion (MacGarvie, 2006; Maggioni et al., 2007; Miguelez, 2016; Miguelez & Noumedem, 2017).

Our analysis takes place at the country pair level, as follows:

$$KC_{ijt} = e^{\alpha_0}.NOHSShare_{ijkt}^{\alpha_1}.Z_{ijt}^{\gamma_n}.e^{\tau_i}.e^{\tau_j}.e \qquad (1)$$

Where

- $KC_{ijt}$ stands alternatively for one of our two dependent variables, namely co-inventorship and co-authorship between OECD countries *i* and *j* at time *t*[2], and;
- $NOHSShare_{ijkt}$ is the product of the share in percentage of *hs* migrants from non-OECD country *k* residing respectively in countries *i* and *j*; (notice that all our country pairs consist of destination countries, with migrants coming from outside these destination countries)
- $Z_{ijt}$ is a set of *n* dyadic covariates and country specific control variables at time *t;*
- $\tau_i$, $\tau_j$ and $\delta_t$ are country *i*, country *j* and time fixed effect (FE) respectively*;*
- $\varepsilon_{ijt}$ is the error term.

[2] In the cross-section regressions we run, the time *t* corresponds to a five-year window: 2000 – 2004 and 2010 – 2014. Hence we do a five-year average of all variables except for time invariant variables and the *hs* migrants variable.

Both in the trade or migration literature, due to the presence of many zeros in the dependent variable the multiplicative form of the gravity equation is often transformed into its logarithmic form before being regressed with an Ordinary Least Square (OLS). However, this may induce heteroskedasticity in the error terms, hence some inconsistency of estimation. As a remedy, we follow (Santos-Silva & Tenreyro, 2006) methodology and estimate the gravity equation using a Poisson pseudo-maximum likelihood (PPML) model. One advantage of PPML estimations is that when the dependent variable is a count variable and the covariates are in logarithms, coefficient estimates can be interpreted as elasticities. We therefore transform equation (1) into its conditional expectation form as follows:

$$E(KC_{ijt}|X_{ijt}) = \exp[\alpha_0 + \alpha_1 \ln NOHSShare_{ijkt} + \gamma_n \ln Z_{ijt} + \tau_i + \tau_j + \delta_t + \quad (2)$$

We add a unit to all explanatory variables in order to correct for the zero values in the natural logarithmic transformation. Following (Felbermayr et al., 2010), we run cross-section regressions for each period *t*.

In further model specifications, we introduce the *OHS* migrant bilateral links variable – $OHSBShare_{ijkt}$ – in the baseline model in order to control for the migration effects from *OHS* migrants themselves. This variable is the product of the percentage of *hs* migrants from country *i* living in country *j* and the percentage of *hs* migrants from country *j* living in country

### *4.2 Data*

We consider as destinations 31 OECD member countries (see Table A1 in Appendix A) out of 35, as of 2017[3]. They account for the largest share of cross-border flows of R&D worldwide (De Backer & Basri, 2008) and include the top five *hs* migrants host countries in the world. As for the origin countries, we focus on India and China, which are the top non-OECD countries of origin for *hs* migrants (Widmaier & Dumont, 2011) (see also Table A6 in Appendix A)[4]. Additionally, India and China are among the fastest growing migrant countries of origin.

[3] We exclude Turkey due to some data inconsistency in foreign migration records. We also exclude South Korea, Latvia and Iceland, for which we have no or severely incomplete migration figures.

[4] Indeed, the other top sending countries are largely EU countries; which generates a mix of intra-regional – intra-EU – and international – EU-US or EU-Canada – migration patterns.

Our observations are country pairs. For the dependent variable – cross-country scientific and technology collaboration – we consider two alternative proxies: co-inventorship and co-authorship. Each of them captures different features of the phenomenon of interest. We examine them in turn.

### 4.2.1 Dependent variables

We define co-inventorship by considering inventors collaboration on a single patent. This helps capturing the overall joint inventorship between countries, regardless of the different standing of partners. This implies that co-inventors may or may not come from the same company – within Multinational Corporations (MNCs). Therefore although computed at the individual level, the co-inventorship variable may point to links between individuals, organizations or individual-organization. The latter is one important feature which differentiates the co-inventorship variable from the co-authorship one as it will be seen below.

We compute the co-inventorship variable from raw data. In particular, we use the February 2015 version of the OECD REGPAT database, which covers patent applications to the European Patent Office (EPO) – as derived from PatStat, the Worldwide Patent Statistical Database released in Autumn 2014 – and PCT patents – the patents database from the World International Patents Office (WIPO) (OECD, REGPAT database, Feb., 2015)[5]. This database records information on patent inventors and their addresses. We check for duplicates on patents application in both PCT and EPO data on the basis of their application number, applicant names and addresses and remove them. For the international co-inventorship variable we only keep patent records with at least two co-inventors with their respective addresses in two distinct OECD countries. Then we proceed to counting co-inventorships by pair of countries, by year. Each patent generates as many co-inventorship counts as the number of combinations of OECD countries we can obtain from the inventors' addresses. The more inventors from two OECD countries are listed on the same patent, the higher the count. We do not record the combinations involving one or two non-OECD countries. Tables 1a and 1b provide an example: patent WO2005100777 reports three inventors from Germany and as many from Austria, which generates 9 Germany-Austria co-inventorships ; patent EP1320536 reports three inventors from Switzerland and one from France, which generates 6 Switzerland-France co-inventorships ; and so forth.

[5]See (Squicciarini et al., 2013) for a detailed description of the database.

**Table 1a:** Example of co-inventorship data (two patents)

| Patent id | Inventors' Country | Year |
|---|---|---|
| WO2005100777 | AT | 2004 |
| WO2005100777 | AT | 2004 |
| WO2005100777 | AT | 2004 |
| WO2005100777 | DE | 2004 |
| WO2005100777 | DE | 2004 |
| WO2005100777 | DE | 2004 |
| WO2005100777 | IT | 2004 |
| WO2005100777 | IT | 2004 |
| EP1320536 | CH | 2000 |
| EP1320536 | CH | 2000 |
| EP1320536 | CH | 2000 |
| EP1320536 | FR | 2000 |
| EP1320536 | GB | 2000 |

**Table 1b**: Count of co-inventorships, by country pairs from data in Table 1a

| Ctry1 | Ctry2 | Co-inventorships | Year |
|---|---|---|---|
| AT | DE | 9 | 2004 |
| AT | IT | 6 | 2004 |
| DE | IT | 6 | 2004 |
| CH | FR | 3 | 2000 |
| CH | GB | 3 | 2000 |
| FR | GB | 1 | 2000 |

Our co-authorship variable measures the extent of collaboration for scientific or basic knowledge production purposes. Publication data cover a wider range of fields than patent data. Hence, the co-authorship variable stands as a broader scope proxy for collaborations. Additionally the co-authorship variable fully captures inter-institutions collaboration, insofar that international collaborations recorded in publications are commonly made between authors affiliated to different institutions; which is not the case in patents where intra-company collaborations are more frequent. The co-authorship data were provided by the OECD, based on elaborations of Scopus Custom Data[6].

[6]That is a customised large-scale dataset for research performance analysis derived from Scopus core records. Scopus itself is a bibliometric database owned by Elsevier which contains abstracts and references from over 22,748 peer-reviewed journals from 4,000 publishers worldwide; thus a wider multidisciplinary coverage.

The co-authorship variable is built at the country-pair level, based on information on authors' country of affiliation as recorded on publication. An affiliation can be a university or any research institution. Thus the co-authorship variable is just the sum of co-authorships per country pair, per year. Similarly to co-inventorship, we count each co-authored publication across each country pair as single instance – one unit – of co-authorship, no matter the number of authors involved. We only consider co-authorships among the OECD countries of our sample.

### 4.2.2 Explanatory variables and controls

In order to track the *NOHS* diasporas, we need information on *hs* migrants from China and India to OECD destination countries. Additionally, we need data on the bilateral *hs* migration between OECD countries so to compute the *OHS* migrant bilateral links variable.

Hence, we use the first and third editions of the OECD-DIOC database (Database on Immigrants in OECD Countries; 2000/01, 2010/11)[7], which assembles information from various national sources on the stock of immigrants for each of the OECD destination countries, from around 200 origin countries. Table A1 in Appendix A provides details on the data source for each destination country.

Immigrants are mostly identified on the basis of their place of birth[8]. Additionally, DIOC reports information on the migrants skill level, as proxied by their educational attainments. We consider as *highly skilled* all the individuals who have completed a tertiary level of education; in other words, those belonging to levels 5 and 6 according to the International Standard Classification of Education (ISCED) of UNESCO. We compute the *NOHS* diaspora variable ($NOHSShare_{ijkt}$) as the product of the percentage of *hs* migrants from India/China in the total *hs* population, for each pair of OECD receiving countries. As for the *OHS* migrant bilateral links variable ($OHSBShare_{ijt}$) for each OECD country pair we compute the product of the percentage of *hs* migrants from one country and residing in the other country of the pair.

[7]There are three editions of DIOC: 2000/01, 2005/06 and 2010/11, which are demographic data collected on 34 OECD countries based on national censuses, population registers and labour force surveys conducted over a given period approximately throughout the sample of destination countries – between 1999 and 2003 for DIOC 00/01 and 2010 to 2013 for DIOC 10/11. In parallel, there is an extension of DIOC 00/01, DIOC-E which includes 66 additional non-OECD destination countries, We only work with DIOC 00/01 and 10/11 as in the DIOC 05/06 data for many receiving countries like UK, information on immigrants' origin has been aggregated at the regional level, which makes it impossible to identify immigrants from Indian or Chinese origin. It is therefore not possible to get figures on Indian and Chinese hs migrants in these OECD destination countries for the period covered by DIOC 05/06.

[8]Except for Germany and Japan where only information on nationality is available.

The *NOHS* diaspora variable, our focal explanatory variable, is computed as below:

$$NOHSShare_{\mathrm{ijkt}} = \frac{H_{\mathrm{ik}}}{H} \quad (3)$$

where $H_{ikt}$ and $H_{jkt}$ are the *hs* migrants population from country *k* – with *k* equals either to China or India – respectively in OECD countries *i* and *j* at time t, while $H_{it}$ and $H_{jt}$ are the total *hs* population in the same countries at the same time.

The *OHS* migrant bilateral links variable is computed as followed:

$$OHSBShare_{\mathrm{ijt}} = \frac{H}{I} \quad (4)$$

where $H_{ijt}$ and $H_{jit}$ are respectively the total *hs* migrant populations from countries *i* in at time t, and vice versa.

Our main controls consist of a set of dyadic and country-specific variables that account for the effect of other factors affecting the intensity S&T collaboration.

At a dyadic level, we control for the physical and cultural distance between country pairs with several variables from the *'Centre d'Etudes Prospectives et d'Informations Internationales'* (CEPII)[9]. First, we consider two variables for physical distance. One is a dummy variable for contiguity, taking value 1 if the two countries in the observation share a common border and 0 otherwise. The other one is a variable measuring the distance in kilometers between the biggest cities of both countries, weighted by the share of these cities' population in the overall country's population. Second, cultural ties are proxied with a dummy for common language taking value 1 if a language is spoken by at least 9% of the population in both countries and 0 otherwise. Last, for historical ties we include a dummy taking value 1if there has been a past colonial link between both countries and 0 otherwise. Additionally in the co-inventorship equation, we include an index of technological proximity that controls for the similarity of technological specializations of the two countries. This index is computed using patents data from the CRIOS–Patstat database, which is derived from the 2014 EPO patent database (Coffano & Tarasconi, 2014). In particular:

[9]A detailed description of these variables can be found in (Mayer & Zignago, 2011).

$$Tech.similarity_{ij} = \frac{\sum f_{ih} f_{jh}}{(\sum f_{ih}^2 \sum f_{jh}^2)^{1/2}} \qquad (5)$$

where $f_{ih}$ stands for the share of patents of technological class *h* – according to 30-class reclassification of IPC codes[10] - of country *i*, and $f_{jh}$ the share of patents of technological class *h* of country *j*. Values of the index close to the unity indicate that countries of a given pair are technologically similar, and values close to zero mean they are technologically far from each other (Jaffe, 1986).

At the country level, we add technological masses – which in the gravity model help testing for attraction level between both countries – that control for countries specific characteristics in terms of science and technology intensity or capability. For the co-inventorship equation, we interact the five-year averages for each of the three waves – 2000-04, 2005-09, 2010-14 – of the total patents in countries *i* and *j*. That is, we calculate the product of the five-year average of *total patents* in the two countries. The latter figure comes from the 2015 WIPO statistics database. It is the total number of patents per applicants origin filed at the European Patent Office (EPO)[11]. Similarly, we add an interaction term for the five-years averages of publications in countries *i* and *j* in the co-authorship equation. These total publications are derived from the SCImago Journal & Country Rank[12] – see SCImago, 2007 –, which we obtained from the Scopus database.

Last but not the least, we control for economic masses by adding the product of the five-year averages of GDPs for the countries in each pair of countries *i* and *j*.

[10]This 30-class reclassication of IPC codes was originally proposed by the OST (*Observatoire des Sciences et Techniques*). For more details see (Coffano & Tarasconi, 2014).

[11]Using this data, one might think of a potential bias occurring for an over-representation of European patents in the EPO – like Germany for instance – as compared with patents from other parts of the world. However, we doubt this actually biases our results for several reasons. First most OECD countries are European, and for the rest of countries, except for the US, Japan and Israel their inventive activities is comparatively insignificant. Secondly, assuming we have overestimated the technological mass of European countries, the results from table 6 show the potetial over estimation might not be high. If we look at the patents product variable estimates, in all model specifications without the US (as one of the assumed underestimated weight), this variable coefficient drops drastically (sometimes by over half). This implies the weight of the US is still very much important. As for Japan, there aren't many co-inventorhip instances with that country.

[12]SCImago Journal & Country Rank is an online portal that reports journals and country scientific indicators developed from information recorded in SCOPUS. Both data sources are comprehensive at the geographical and thematic levels in a sense that they cover a large range of countries worldwide and a wide interdisciplinary content.

**Tables 2** Variables definition

| Variables | Definition |
|---|---|
| | ***Dependent variable*** |
| Co-inventorship | Count of the number of joint patents inventorship instances between countries *i* and *j*. |
| Co-authorship | Count of the number of joint publications authorship instances between countries *i* and *j*. |
| | ***Explanatory and control variables*** |
| *NOHS* diaspora | Product of the percentage of *NOHS migrants* from same origin in countries of destination *i* and *j*. In other words, it is the probability that if we select an individual at random from each of the two countries' total *hs* population, both will be from the same *NOHS* migrant country of origin. |
| *OHS* migrant bilateral links | Product of the percentage of *OHS* migrants from country *i* living in country *j* and the percentage of *OHS* migrants from country *j* living in country *i*. Or it is the probability that, if we select an individual at random from each of the two destination countries total *hs* population, both will be*hs* migrants from one of the two countries living in the other. |
| Technological similarity | Index of technological proximity between countries *i* and *j* varying from 0 to 1. |
| Contiguity | Dummy variable for taking value 1 if countries *i* and *j* share same border and 0 otherwise. |
| Colony | Dummy variable taking value 1 if both countries *i* and *j* share a common past colonial history and 0 otherwise. |
| Common official langage | Dummy variable taking value 1 if same language is spoken by at least 9% of the population in countries *i* and *j* |
| Distance | Distance in kilometres between the biggest cities of both countries *I* and *j*, weighted by the share of these cities' population in the overall countries' population |
| Product of patents in countries *i* and *j* | Product of the total # of patents in countries *i* and *j*; $\#Pat_i*\#Pat_j$ |
| Product of publications in countries *i* and *j* | Product of the total # of publications in countries *i* and *j*; $\#Pub_i*\#Pub_j$ |
| Product of GDP in countries *i* and *j* | Product of the GDP in countries *i* and *j*; $\#GDP_i*\#GDP_j$ |

### 4.3 Descriptive statistics

Based on the 31 OECD receiving countries in our sample, we have 465 observations per DIOC edition (DIOC 00/01 and 10/11) and per immigrants' country of origin (Chine and India).

Table 3 below depicts the top fifteen S&T collaboration corridors between 2010 and 2014. More precisely, it shows the five-years averages of co-inventorship and co-authorship for the fifteen country pairs with the highest figures.

One important evidence from the table is the leading position of the US as a key research partner for many European and non-European countries. Indeed, the country is present in ten country-pair collaborations, in both patenting and publication out of the top fifteen collaborations of our table.

**Table 2:** Science and technology research corridors for years 2010 – 2014

| Country a | Country b | Average co-inventorship* | Country a | Country b | Average co-authorship* |
|---|---|---|---|---|---|
| USA | Germany | 3,388 | USA | UK | 22,334 |
| USA | UK | 2,629 | USA | Germany | 19,865 |
| USA | Canada | 2.544 | USA | Canada | 18,687 |
| Germany | Switzerland | 1,572 | USA | France | 13,013 |
| France | Germany | 1,397 | UK | Germany | 11,693 |
| USA | France | 1,201 | USA | Italy | 11,366 |
| USA | Japan | 1,171 | USA | Japan | 10,116 |
| USA | Switzerland | 927 | USA | Australia | 9,898 |
| USA | Israel | 804 | France | Germany | 8,615 |
| Germany | Austria | 717 | UK | France | 8,430 |
| UK | Germany | 679 | USA | Spain | 8,353 |
| France | Switzerland | 641 | USA | Netherlands | 7,930 |
| USA | Netherlands | 604 | Italy | UK | 7,514 |
| USA | Belgium | 571 | Germany | Switzerland | 7,221 |
| USA | Italy | 568 | USA | Switzerland | 7,180 |

* This number corresponds to the five-year average of S&T collaborations between 2010 and 2014

The Chinese *hs* migrant figures from DIOC 10/11 and DIOC 00/01 are reported in Table 4 below. The US stand out as the country with the biggest intake of Chinese *hs* in absolute terms in both DIOC editions; from a total number of over 372,000 Chinese *hs* in the first wave to over 484,000 Chinese *hs* in the third edition, so an increase of nearly 30%. However, this increase is smaller than that of all other receiving countries, where the total number of Chinese *hs* has more than doubled or tripled over the same period of time. For instance in Canada – which is the second top receiving country of the list – the intake of Chinese *hs* has gone from over 86,000 to 208,000, an increase of almost 142%. A glance at the columns reporting the shares of Chinese *hs* on the total *hs* population in each receiving country shows a similar story. The share of Chinese *hs* migrants over total *hs* from DIOC 00/01 to DIOC 10/11 has increased in all receiving countries. Interestingly, Australia appears as the country with the biggest shares of Chinese *hs* in the total *hs* population with 1.43% and 3.13% in the first and third editions respectively.

**Table 3**: Top fifteen receiving countries of Chinese *hs* migrants in DIOC 10/11 and 00/01

| DIOC 10/11 | | | | DIOC 00/01 | | | |
|---|---|---|---|---|---|---|---|
| **Country of destination** | **Chinese *hs* immigrants ('000)** | **Total *hs* residents ('000)** | **Chinese share of *hs* (%)** | **Country of destination** | **Chinese *hs* immigrants ('000)** | **Total *hs* residents ('000)** | **Chinese share of *hs* (%)** |
| USA | 484.2 | 59,088 | 0.82 | USA | 372.8 | 46,304 | 0.81 |
| Canada | 208 | 9,155 | 2.27 | Canada | 86 | 6,320 | 1.36 |
| Australia | 116.2 | 3,713 | 3.13 | Japan | 43.9 | 21,125 | 0.21 |
| UK | 76.5 | 11,827 | 0.65 | Australia | 32 | 2,252 | 1.43 |
| Japan* | 75.1 | -------- | ---- | UK | 11.7 | 6,856 | 0.17 |
| New Zealand | 25.1 | 1,018 | 2.47 | Germany | 6.2 | 8,540 | 0.07 |
| Germany | 22.9 | 11,610 | 0.20 | France | 5.3 | 6,305 | 0.08 |
| France | 16.9 | 9,717 | 0.17 | New Zealand | 5 | 567 | 0.88 |
| Spain | 7 | 8,514 | 0.08 | Sweden | 1.8 | 1,137 | 0.16 |
| Italy | 4.6 | 4,797 | 0.10 | Switzerland | 1.7 | 824 | 0.20 |
| Sweden | 4.6 | 1,575 | 0.29 | Spain | 1.4 | 4,266 | 0.03 |
| Switzerland | 4.2 | 1,439 | 0.29 | Italy | 1.3 | 3,004 | 0.05 |
| Netherlands | 3.6 | 2,763 | 0.13 | Belgium | 1 | 1,432 | 0.07 |
| Ireland | 2.8 | 786 | 0.35 | Austria | 0.7 | 582 | 0.12 |
| Austria | 2.0 | 809 | 0.25 | Ireland | 0.6 | 562 | 0.11 |

*For Japan we were not able to compute total *hs* figures from the DIOC 10/11 data as all education levels for Japanese nationals were recorded as unknown.

The figures for the top Indian *hs* receiving countries are very similar to those for the Chinese, except for a few cases. In Table 5, the US appear to be the top receiving country of Indian *hs* in absolute terms in both DIOC editions – from over 504,000 to nearly 975,400. In DIOC 10/11 figures, Canada loses to the UK its second position occupied in DIOC 00/01 – from a total Indian *hs* of over 101,100 to

230,400 for Canada and from 96,900 to 269,400 for the UK. As for the shares of Indian *hs* in the total *hs* population, they have increased in all countries from the first to the third DIOC edition. Canada and Australia are the countries with the highest shares in the first and third editions respectively – 1.60% and 4% respectively.

**Table 4**: Top fifteen receiving countries of Indian *hs* migrants in DIOC 10/11 and 00/01

| DIOC 10/11 | | | | DIOC00/01 | | | |
|---|---|---|---|---|---|---|---|
| **Country of destination** | **Indian *hs* *immigrants* *('000)*** | **Total *hs* *residents* *('000)*** | **Indian share *hs* (in%)** | **Country of destination** | **Indian *hs* *Immigrans* *('000)*** | **Total *hs* *residents* *('000)*** | **Indian share *hs* (in%)** |
| USA | 975.4 | 59,088 | 1.65 | USA | 504 | 46,304 | 1.09 |
| UK | 269.4 | 11,827 | 2.28 | Canada | 101.1 | 6,320 | 1.60 |
| Canada | 230.4 | 9,155 | 2.52 | UK | 96.9 | 6,856 | 1.41 |
| Australia | 148.5 | 3,713 | 4.00 | Australia | 34 | 2,252 | 1.51 |
| New Zealand | 30.7 | 1,018 | 3.01 | New Zealand | 5.8 | 567 | 1.02 |
| Germany | 9.9 | 11,610 | 0.09 | France | 3.9 | 6,305 | 0.06 |
| Ireland | 9.8 | 786 | 1.25 | Israel | 2.4 | 1,036 | 0.23 |
| France | 8.4 | 9,717 | 0.09 | Switzerland | 2.4 | 824 | 0.29 |
| Switzerland | 7.2 | 1,439 | 0.50 | Japan | 1.9 | 21 | 0.01 |
| Italy | 5.6 | 4,797 | 0.12 | Netherlands | 1.9 | 1,911 | 0.10 |
| Sweden | 4 | 1,575 | 0.25 | Sweden | 1.7 | 1,137 | 0.14 |
| Israel | 3.7 | 1,531 | 0.24 | Italy | 1.6 | 3,004 | 0.05 |
| Netherlands | 3.1 | 2,763 | 0.11 | Ireland | 1.5 | 562 | 0.26 |
| Spain | 3 | 8,514 | 0.04 | Belgium | 1.5 | 1,432 | 0.10 |
| Japan* | 2,9 | --------- | ---- | Norway | 1 | 640 | 0.17 |

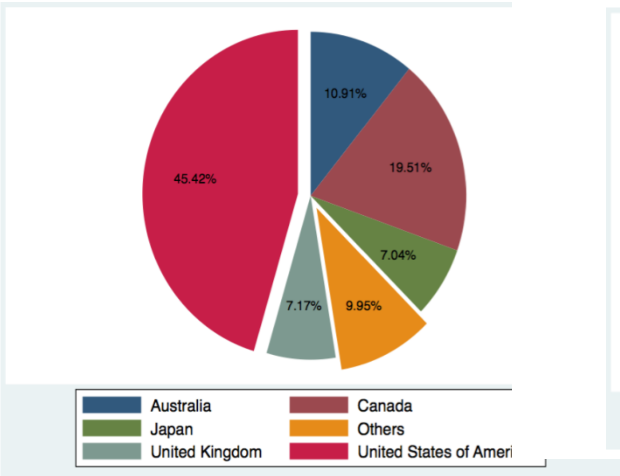
*For Japan we were not able to compute total *hs* figures from the DIOC 10/11 data as all education levels for Japanese nationals were recorded as unknown.

Overall, these figures show a skewed distribution of Indian and Chinese *hs* migrants across countries, as these migrants concentrated mostly in the top four countries of the list. These distributions are better illustrated in the following graphs.

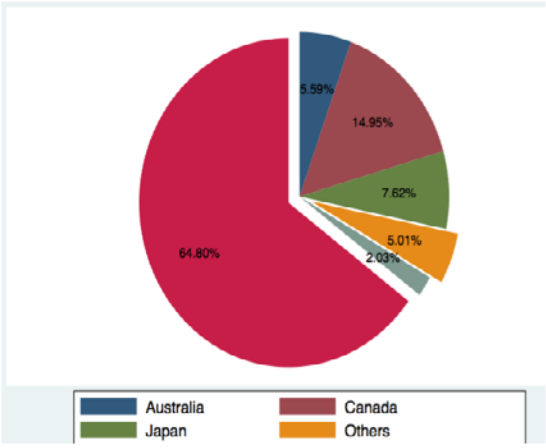
In Figure 2, we see the change in the distribution of Chinese *hs* migrants in DIOC 10/11 and DIOC 00/01 respectively for the OECD countries of our sample. These graphs show the biggest intakes of Chinese *hs* to be in five countries only, out of the 31 countries in our sample. These countries are the US, Canada, Japan, Australia and the United Kingdom, whose total share of Chinese *hs* accounts for around 90.05% and 94.99% of the total Chinese *hs* migrants in our sample, in DIOC 10/11 and DIOC 00/01 respectively. However it is important to mention the differences between the two graphs, with the US losing its shares of Chinese hs – from 64.80%, to 45.42% - to the benefit of other destination countries. In other words, Chinese *hs* migrants seem to find other destination countries such as Australia and the UK increasingly more attractive. This might be due to factors such as the loosening of emigration controls in China (Ortega & Peri, 2009), but also the gradual shift of immigration policies in some European countries like in France, away from their traditional focus on family reunions and asylum seeking to more *hs*-oriented policies (Docquier et al., 2007).

**Figure 2**: Chinese *hs* migrants distribution in 31-OECD destination countries: DIOC 10/11 vs. DIOC 00/01

DIOC 10/11



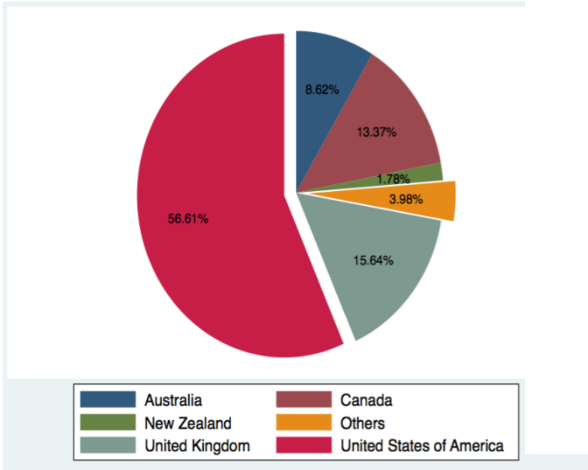

DIOC 00/01



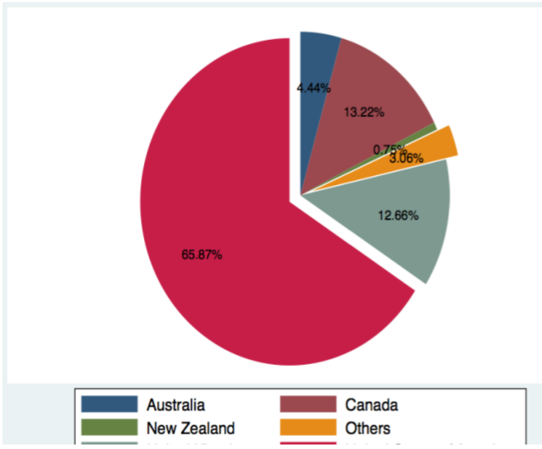

Similar remarks can be made from the figures depicting the distribution of Indian *hs* migrants in our sample of receiving countries as shown in Figure 3 below. The distribution tends to be slowly becoming more even from DIOC 00/01 on the right hand side to DIOC 10/11 on the left. But this process is seemingly confined to the group of the four biggest receiving countries; the US, Canada, the United Kingdom and Australia. Hence, this implies some clustering effect of the Indian *hs* diaspora in these countries. Interestingly, an important share of these Indian *hs* diaspora is made of skilled professionals, business scientists and academic elites (Docquier & Rapoport, 2012), and these countries seem to be more attractive to these countries.

**Figure 3**: Indian *hs* migrants distribution in DIOC 10/11 and DIOC 00/01

DIOC 10/11




DIOC 00/01




## 5 Results

### 5.1 The effect of the Chinese and Indian hs diaspora in S&T collaborations

Table 6 below reports the results from the co-inventorship regressions for Chinese and Indian *hs* diasporas for DIOC 10/11 and DIOC 00/01 respectively, with country and time fixed effect. The coefficient estimates are elasticities. The results for the Chinese *hs* diaspora are illustrated in columns (1) to (6) – with columns (1) to (3) showing DIOC 00/01 results and columns (4) to (6) showing those from DIOC 10/11 -, while columns (7) to (12) present the outcomes from the Indian *hs* diaspora in DIOC 00/01 - (7) to (9) – and DIOC 10/11 respectively - (10) to (12). For each *NOHS* diaspora variable and for each DIOC edition, we use three model specifications.

In the baseline model we only include our main variable of interest, $lnNOHSShare_{ijkt}$, along with basic control variables. We then add the *OHS* migrant bilateral links variable $lnOHSBShare_{ijt}$ to control for the effect of *hs* migrants exchanges between host countries *i* and *j*. Last, as a robustness check, we drop from our sample all country pairs that include the US– so to check whether it is that specific country which drives our results, due to its double status as both the top *hs* migrants host country and the most recurrent partner in international R&D collaborations, within and across firms.

From the baseline regressions for the Chinese *hs* variable – columns (1) and (4) –, we get different results for the elasticities of our main variable of interest in DIOC 00/01 and DIOC 10/11 respectively. Indeed, while for the first DIOC edition we get a strong positive and statistically significant result – 1.321 –, for the second edition the coefficient value decreases to 0.285. In other words we find that, if we double the probability of getting Chinese migrants from a random draw of two individuals from two host countries total *hs* population, co-inventorship between these two countries will increase by 132.1% in DIOC 00/01 and by 28.5% in DIOC 10/11. Projecting these results into our sample average figures (Table A2 in Appendix A) gives us the marginal effects. Our sample average probability of having two Chinese *hs* in a country-pair random draw is 0.041 in DIOC 00/01. If that number doubles to0.082 the induced effect will be an important raise in the average number of co-inventorships among country pairs to the value of 223[13]. By analogy, if we double the DIOC 10/11 sample average probability of Chinese *hs* from 0.987 to 1.974, this will induce a change in the average number of country pair co-inventorships from 72 to almost 93[14]. Controlling for the *OHS* migrant bilateral links variable $lnOHSBShare_{ijt}$ in the baseline equations only modestly changes the results. As shown in columns (2) and (5) for the first and third DIOC editions respectively, our main coefficient remains positive and statistically significant but slightly drops to 1.284 and 0.269 in DIOC 00/01 and 10/11 respectively. In parallel we get positive and significant results for the *OHS* migrant bilateral links variable - 0.991 and 0.758 respectively. Interestingly, the coefficient for the Chinese *hs* variable is stronger than the one for the *OHS* migrant bilateral links in the DIOC 00/01. When dropping the US as a destination country from our regressions, the magnitude of our main variable coefficient decreases to 0.950and 0.218 in first and second DIOC editions as shown in columns (3) and (6) respectively. As for the coefficients of the usual gravity covariates, they are in line with findings from the trade or migration literature except for a few cases where we get insignificant results.

[13]Tables A2 in the Appendix reports the average number of country-pair co-inventorships in DIOC 00/01 for our sample to be 96. Applying a 132.1% increase to that number we get 127 + 96 which gives 223. In general, we use this rule to get the marginal effect of each of our explanatory variables based on our sample mean values for each of the dependent variables. That is, we get the marginal effect in terms of each of our dependent variables as N(α + 1), where N is the average value of that variable in our sample and α is the coefficient estimate of the migration variable

[14] It is important to note that this number represents the change in the sample average co-inventorship in absolute terms and not the change in the average number of co-patenting. Indeed, our co-inventorship was built by counting all occurrences of country pair inventors' residency in a single patent. Therefore, this change might as well capture an increase of co-inventorship within patents; or an increase of patents co-invented across a country-pair.

Two expected results to stress are the positive and significant coefficients of the variable for the product of total patents in each of the two countries and the variable for technological similarity in all model specifications. The positive and significant sign of the total patents product variable suggests that countries that file more patents are more likely to get involved in co-inventorship activities. While the result of technological similarity suggests that countries performing in technologically similar fields have a greater chance to engage in joint inventorship.

The outcomes for India as an origin country are similar to those for China. The baseline model of DIOC 00/01 and 10/11 in columns (7) and (10) respectively return positive and significant results for the coefficients of our main variable of interest – 1.008and 0.298. This means, if the probability of getting two Indian *hs* from a random draw of two individuals from the total *hs* population of the two host countries doubles, the co-inventorship between these countries would increase by 100.8% in DIOC 00/01, and by 29.8% in DIOC 10/11.Following a similar reasoning as earlier, we translate these results into their marginal effect values by projecting them into our sample average figures from Table A2 in Appendix A. Doubling the DIOC 00/01 sample average probability of having two Indian *hs* in a country-pair random draw from 0.071 to 0.142 would induce an increase in the average co-inventorships from 96 to nearly 193. Similarly, in DIOC 10/11 an increase in the average probability of Indian *hs* from 0.331 to 0.662 leads to an increase in the co-inventorships sample mean from 72 to 94. When adding the *OHS* migrant bilateral links variable to the baseline equation, our main coefficient remains positive and significant as illustrated in columns (8) and (11) respectively – 1.028 and 0.605. Removing the US as a destination country from the data doesn't affect our results sign. We still get positive and significant coefficients, although their value diminishes – 0.448 and 0.249 as shown in columns (9) and (12) respectively. This suggests Indian *hs* diaspora seem to have a major impact on co-inventorship activities between the US and the rest of OECD countries. Other control variable results do not differ much from what we had for the case of China as described above. Again here we get positive and significant results for the estimates of the product of total patents and technological similarity index.

**Table 5**: Chinese and Indian *hs* diaspora in international co-inventorship

| *Co-inventorship* | DIOC 00/01 CHINA | | | DIOC 10/11 CHINA | | | DIOC 00/01 INDIA | | | DIOC 10/11 INDIA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *$lnNOHSShare_{ijk}$* | 1.321** | 1.284** | 0.950** | 0.285** | 0.269*** | 0.218** | 1.008*** | 1.028*** | 0.448* | 0.298* | 0.605*** | 0.249** |
| | (0.638) | (0.637) | (0.401) | (0.126) | (0.093) | (0.102) | (0.273) | (0.283) | (0.260) | (0.163) | (0.186) | (0.118) |
| *$lnOHSBShare_{ij}$* | | 0.991*** | 0.694*** | | 0.758*** | 0.650*** | | 0.937*** | 0.461*** | | 0.745*** | 0.550** |
| | | (0.218) | (0.179) | | (0.148) | (0.219) | | (0.286) | (0.148) | | (0.286) | (0.227) |
| *Common lang.* | 0.558*** | 0.344** | 0.467*** | 0.469*** | 0.377*** | 0.434*** | 0.512** | 0.346 | 0.476** | 0.609*** | 0.499*** | 0.361*** |
| | (0.196) | (0.170) | (0.163) | (0.099) | (0.083) | (0.136) | (0.235) | (0.211) | (0.192) | (0.191) | (0.178) | (0.1228) |
| *ln(distance)* | -0.346*** | -0.403*** | -0.371*** | -0.117 | -0.095 | 0.093 | -0.322*** | -0.348*** | -0.287*** | -0.329** | -0.239** | -0.337*** |
| | (0.109) | (0.104) | (0.070) | (0.094) | (0.071) | (0.080) | (0.111) | (0.106) | (0.084) | (0.097) | (0.094) | (0.057) |
| *Contiguity* | 0.221 | -0.082 | 0.315** | 0.593*** | 0.398*** | 0.465*** | 0.274 | 0.024 | 0.310** | 0.303** | 0.108 | 0.350*** |
| | (0.171) | (0.196) | (0.128) | (0.128) | (0.122) | (0.125) | (0.181) | (0.215) | (0.139) | (0.135) | (0.169) | (0.122) |
| *Colony* | 0.431** | 0.325* | 0.166 | -0.072 | -0.148 | 0.225 | 0.170 | 0.045 | 0.022 | 0.205 | 0.085 | -0.217 |
| | (0.201) | (0.179) | (0.202) | (0.100) | (0.090) | (0.149) | (0.206) | (0.220) | (0.235) | (0.168) | (0.190) | (0.138) |
| *Tech. similarity* | 3.400* | 2.902* | 2.057* | 3.193*** | 3.254*** | 2.736*** | 3.413* | 3.176* | 1.443 | 3.103** | 3.011** | 3.074*** |
| | (1.806) | (1.715) | (1.158) | (0.717) | (0.620) | (0.601) | (2.022) | (1.880) | (1.248) | (1.490) | (1.472) | (0.643) |
| *$ln(GDP_i*GDP_j)$* | 1.296** | 1.524*** | 0.412 | 0.615 | -0.619 | 0.830*** | 1.745*** | 1.834*** | 0.243 | 2.033*** | 2.340*** | 0.830*** |
| | (0.590) | (1.505) | (0.309) | (1.267) | (0.489) | (0.310) | (0.654) | (0.616) | (0.219) | (0.543) | (0.615) | (0.230) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *ln(patent$_i$*patent$_j$)* | 1.094*** | 1.160*** | 0.640*** | 1.110*** | 0.985*** | 0.203*** | 0.759*** | 0.796*** | 0.520*** | 0.850*** | 0.803*** | 0.458*** |
| | (0.152) | (0.143) | (0.093) | (0.175) | (0.063) | (0.059) | (0.088) | (0.080) | (0.069) | (0.061) | (0.047) | (0.041) |
| *Constant* | -39.26*** | -44.25*** | 3.669 | -30.60 | -2.602 | -19.96*** | 43.30*** | -45.57*** | 8.124*** | -53.19*** | -60.78*** | -21.33*** |
| | (13.28) | (11.25) | (6.119) | (29.87) | (10.68) | (6.649) | (13.15) | (12.56) | (4.004) | (11.93) | (13.51) | (5.072) |
| *Observations* | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 |
| *Pseudo R2* | 0.943 | 0.964 | 0.959 | 0.987 | 0.993 | 0.976 | 0.925 | 0.945 | 0.965 | 0.931 | 0.947 | 0.976 |
| *Countries & time FE* | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *Without the US* | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We add a unit to all of the above explanatory variables before logarithmic transformation in order to account of the presence of many zeros

For the co-authorship dependent variable, we run similar model specifications as in the co-inventorship variable section above. We report the results from these regressions in Table 7 below. First, we look at the effect of the Chinese *hs* diaspora in the first and second DIOC edition – columns (1) to (3) and (4) to (6) respectively. Then we run the same regressions again for the case of the Indian *hs* diaspora in DIOC 00/01 – (7) to (9) – and DIOC 10/11 – (10) to (12). The baseline regressions for both first and second DIOC editions – (1) and (4) –, yield positive and statistically significant estimates for the coefficients of the Chinese *hs* diaspora variable, with a value of 0.292 and 0.221 respectively. This means, we find that if theprobability of having two Chinese *hs* from a random draw of two individuals from the total *hs* population of two host countriesdoubles, this would lead to a 29.2% and a 22.1% increase in joint authorship between these countries in the periods covered by DIOC 00/01 and DIOC 10/11 respectively. In order to obtain the marginal effects, we apply the coefficient estimates to the mean values of our sample. In DIOC 00/01, this means an increase of the sample average co-authorship from 407 to 526 would come from doubling the sample average probability of having two Chinese *hs* in a country-pair random draw from 0.041 to 0.082. Similarly, in DIOC 10/11, doubling the sample average probability of having two Chinese *hs* – from of 0.987 to 1.974 – would increase the sample mean co-authorship from 1,220 to 1,490. Moreover, these estimated coefficients remain positive and significant after controlling for the *OHS* migrant bilateral links $lnOHSBShare_{ijt}$ as shownin columns (2) and (5). In DIOC 00/01 the estimate modestly drops to 0.262, while in DIOC 10/11 that estimate is reduced by nearly half of the baseline value to 0.114. As for the coefficients for the *OHS* migrant bilateral links, we get positive and significant values of 0.566 and 0.406 in each DIOC edition as shown in columns (2) and (5) respectively. When removing all country pairs involving the US, the estimates for our main explanatory variable increase to 0.954 in DIOC 00/01 and almost remain unchanged in DIOC 10/11 with a value of 0.111 – see columns (3) and (6) respectively. We also obtain positive and significant estimates for the coefficient of the *OHS* migrant bilateral links in both DIOC editions.

As for the controls, the coefficients for the usual gravity covariates yield mitigated results in all of the first six columns. In general, we find positive and significant results of the coefficients for the colonial ties, common language and contiguity variables and a negative and significant estimate for the geographical distance variable in the baseline models. However, these effects disappear – except for the coefficient of geographical distance – in the models where we introduce the *OHS* migrant bilateral links variable $lnOHSBShare_{ijt}$. This implies the effect of *OHS* migrant bilateral links on co-authorship suffices to absorb the effects of most of the common gravity covariates. Also, we find a negative and significant coefficient of the variable for the product of GDPs of -0.219 in columns from (4), suggesting countries joining their efforts in co-authorship activities do not always belong to the same income group. Additionally, we find countries level of publications to positively impact on joint authorship, as shown by the positive and significant coefficient for product of publication levels in two countries in all model specifications.

Findings on impact of Indian *hs* migrants on co-authorship are in line with those for the Chinese, but with minor exceptions. The results for our main variable coefficients $lnNOHSShare_{ijkt}$ do not differ much from results for the Chinese *hs* diaspora variable in the baseline model as shown in columns (7) and (11) for DIOC 00/01 and 10/11 respectively – 0.246 and 0.229. We find a 100 percentage point increase in the Indian *hs* migrant diaspora would result to an increase in co-authorship by 24.6% and 22.9% in the periods covered by first and second DIOC editions respectively. Similarly here, we use our sample average values to get the marginal effects. This means that, in DIOC 00/01, doubling the average probability of getting two Indian *hs* from a random draw from two host countries – from 0.071 to 0.142 – would lead to an increase in the average co-authorship from 407 to 507. In a similar fashion in DIOC 10/11, if we double that same average share of two Indian *hs* from 0.331 to 0.662, this would result in an increase in the sample average co-authorship size from 1220 to 1499. When introducing the *OHS* migrant bilateral links variable $lnOHSBShare_{ijt}$ into the baseline model, our regressions yield slightly weaker estimates for the Indian *hs* diaspora variable – 0.217 and 0.114 – in the first and second DIOC editions respectively – see columns (8) and (11). The estimates for the *OHS* migrant bilateral links remain positive and significant in all model specifications. Dropping the US from the regressions leads to an increase of the value of our main variable coefficients, 0.341 and 0.138 in DIOC 00/01 and 10/11 as it can be seen in columns (9) and (12) respectively. Other control variables and common gravity covariates estimates return results which are similar to what we had from the Chinese *hs* diaspora regressions insofar that these results are mitigated in their signs and significance, while we always get positive and significant estimates of the variable for the product of publications size in all columns from (7) to (12).

**Table 6**: Chinese and Indian *hs* diaspora in international co-authorship

| *Co-authorship* | DIOC 00/01 CHINA | | | DIOC 10/11 CHINA | | | DIOC 00/01 INDIA | | | DIOC 10/11 INDIA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *lnNOHSShare*$_{ijk}$ | 0.292* | 0.262* | 0.954*** | 0.221*** | 0.114*** | 0.111*** | 0.246* | 0.217* | 0.341* | 0.229*** | 0.114* | 0.138** |
| | (0.171) | (0.145) | (0.164) | (0.048) | (0.040) | (0.043) | (0.147) | (0.114) | (0.176) | (0.060) | (0.072) | (0.065) |
| *lnOHSBShare*$_{ij}$ | | 0.566*** | 0.407*** | | 0.406*** | 0.292*** | | 0.555*** | 0.366*** | | 0.406*** | 0.274** |
| | | (0.075) | (0.089) | | (0.075) | (0.104) | | (0.116) | (0.133) | | (0.094) | (0.013) |
| *Common lang.* | 0.129* | 0.062 | 0.111 | 0.159*** | 0.141*** | 0.225*** | 0.119 | 0.047 | 0.121 | 0.058 | 0.110 | 0.180** |
| | (0.074) | (0.066) | (0.079) | (0.058) | (0.054) | (0.078) | (0.088) | (0.081) | (0.082) | (0.071) | (0.073) | (0.074) |
| *ln(distance)* | -0.315*** | -0.309*** | -0.323*** | -0.239*** | -0.225*** | -0.226*** | -0.304*** | -0.383*** | -0.319*** | -0.269*** | -0.295*** | -0.266*** |
| | (0.046) | (0.026) | (0.023) | (0.029) | (0.018) | (0.021) | (0.042) | (0.031) | (0.029) | (0.022) | (0.030) | (0.026) |
| *Contiguity* | 0.150* | -0.022 | -0.061 | 0.179*** | 0.077 | 0.071 | 0.153* | -0.102 | 0.055 | 0.140** | -0.001 | 0.045 |
| | (0.080) | (0.070) | ((0.049) | (0.059) | (0.053) | (0.051) | (0.078) | (0.081) | (0.061) | (0.056) | (0.060) | (0.054) |
| *Colony* | 0.149* | 0.030 | 0.224*** | 0.108* | 0.013 | 0.113 | 0.060 | 0.058 | 0.084 | 0.159* | 0.084 | 0.081 |
| | (0.088) | (0.071) | (0.084) | (0.066) | (0.060) | (0.085) | (0.076) | (0.070) | (0.124) | (0.081) | (0.068) | (0.093) |
| *ln(GDP*$_i$**GDP*$_j$*)* | -0.109 | 0.018 | 0.012 | -0.219** | -0.006 | -0.044 | 0.170 | - 0.097* | -0. 134*** | -0.246*** | -0.340** | -0.217*** |
| | (0.140) | (0.069) | (0.066) | (0.089) | (0.070) | (0.078) | (0.110) | (0.052) | (0.052) | (0.068) | (0.928) | (0.086) |
| *ln(puli*$_i$**publi*$_j$*)* | 0.890*** | 0.835*** | 0.761*** | 0.807*** | 0.765*** | 0.744*** | 0.817*** | 0.937*** | 0.752*** | 0.767*** | 0.869*** | 0.645*** |
| | (0.039) | (0.027) | (0.039) | (0.017) | (0.018) | (0.035) | (0.042) | (0.031) | (0.029) | (0.024) | (0.025) | (0.035) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Constant* | -4.638* | -6.293*** | -5.194*** | -1.323 | -5.168*** | -4.019*** | -9.074*** | -5.287*** | -2.103*** | 0.204 | 0.553 | 1.48 |
| | (2.560) | (1.297) | (1.329) | (1.945) | (1.454) | (1.519) | (2.617) | (0.754) | (0.990) | (1.443) | (1.781) | (1.651) |
| *Observations* | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 |
| *Pseudo R2* | 0.973 | 0.982 | 0.979 | 0.976 | 0.983 | 0.977 | 0.972 | 0.975 | 0.974 | 0.976 | 0.975 | 0.977 |
| *Countries & time FE* | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *Without the US* | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We add a unit to all of the above explanatory variables before logarithmic transformation in order to account of the presence of many zeros

### 5.2 Some robustness check

We check for the robustness of our results so far by means of a series of tests conducted in this section and in Appendix B. Here we test for whether our results might be driven by some generic effects from migration, regardless the skill level. Specifically, we test the hypothesis that the positive coefficients for our focal variables we obtained in the previous section could refer to the overall migration patterns, regardless of the skill level of migrants, as indeed is the case with trade. Additionally, this exercise allows us to discuss for potential endogeneity sources in our model – a parallel discussion of endogeneity is conducted in Appendix B with an instrumental variable analysis. In particular, one could argue that the controls we include in our model do not capture entirely the intensity of exchanges between each country pair, so that the estimated coefficients for our focal variables may be affected by a positive omitted variable bias. If this was the case, we would expect to get some positive effect not only from *hs* migration, but also from all of the other skill groups, as the assumed hidden effect would likely affect the entire migration flows between these two countries and not just *hs*.

Hence, we build variables for capturing links connecting migrants in each of the non-OECD low skilled – *ls* – and medium skilled[15]– *ms* – diasporas – $NOLSShare_{ijk}$ and $NOMSShare_{ijk}$ respectively. We then proceed as in the previous section. That is, we calculate the product of the share of *ls* or *ms* Chinese/Indian population in the total *ls* or *ms* population of each country for pairs of countries. Additionally, we build OECD *ls* and *ms* migrant bilateral links variables – $OLSBShare_{ij}$ and $OMSBShare_{ij}$ respectively. We thus run regressions for each of our two dependant variables – co-inventorship and co-authorship–, introducing these new skill variables into our initial equations. The results we get are illustrated in the Table 8 below.

For the co-inventorship regressions, the results show that controlling for Chinese or Indian *ls* and *ms* diasporas actually amplifies the effects of the *hs* variable estimates as their magnitudes significantly increase as compared with what we had in Table 6 above. This increase applies to all model specifications except for column (6) where the estimate of the Indian *hs* diaspora variable is smaller than the one we had in the baseline model. In contrast, we cannot find any significant effect of other skill group variables on co-inventorship. As for the *OHS* migrants' bilateral links variable, its

[15]Low skilled are individuals having a primary education level only while medium skilled are those having a secondary education level.

coefficients also increase considerably from the values we had in the baseline model in Table 6 above.

The co-authorship regressions in DIOC 00/01 return results from which can be drawn similar observations as with the co-inventorship ones: we get an increase of the estimates for the Chinese and Indian *hs* diaspora. However, this is not the case for the estimates in DIOC 10/11 whose effects actually diminish moderately. In parallel, we only get some positive and significant effect for the Indian *ms* diaspora estimates in DIOC 00/01 – see columns (5) and (6). The *ls* network estimates are negative in almost all columns.

**Table 7**: Other Chinese and Indian skill groups and international Co-inventorship and Co-authorship

| | DIOC 00/01 CHINA | | DIOC 10/11 CHINA | | DIOC 00/01 INDIA | | DIOC 10/11 INDIA | |
|---|---|---|---|---|---|---|---|---|
| ***Co-inventorship*** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| *lnNOHSShare$_{ijk}$* | 2.111** | 1.627*** | 0.530*** | 0.398* | 1.077*** | 0.837*** | 0.548*** | 0.708*** |
| | (0.903) | (0.606) | (0.184) | (0.206) | (0.300) | (0.306) | (0.157) | (0.087) |
| *lnNOMSShare$_{ijk}$* | -1.444 | 1.687 | 0.006 | -0.121 | 1.109 | -0.369 | 0.850 | -1.118 |
| | (4.408) | (2.950) | (0.260) | (0.238) | (2.327) | (2.320) | (0.816) | (0.853) |
| *lnNOLSShare$_{ijk}$* | -0.432 | -0.541 | -0.208 | -0.195 | -1.152 | -0.546 | -0.718 | 0.617 |
| | (0.922) | (0.654) | (0.146) | (0.132) | (1.153) | (1.249) | (0.522) | (0.567) |
| *lnOHSBShare$_{ij}$* | | 1.467*** | | 1.256*** | | 1.695** | | 0.934** |
| | | (0.515) | | (0.412) | | (0.664) | | (0.411) |
| *lnOMSBShare$_{ij}$* | | -1.142 | | -1.362 | | -1.531 | | -0.341 |
| | | (1.195) | | (1.155) | | (1.774) | | (1.047) |
| *lnOLSBShare$_{ij}$* | | 0.299 | | 0.424 | | -0.056 | | 0.146 |
| | | (0.588) | | (0.579) | | (0.882) | | (0.568) |
| ***Co-authorship*** | | | | | | | | |
| *lnNOHSShare$_{ijk}$* | *1.439**** | 1.301*** | 0.187* | 0.165* | 0.269* | 0.241* | 0.168** | 0.106* |
| | *(0.532)* | (0.352) | (0.110) | (0.0966) | (0.146) | (0.133) | (0.0741) | (0.0640) |
| *lnNOMSShare$_{ijk}$* | *-1.693* | -1.606 | -0.145 | -0.0788 | 3.015*** | 1.952*** | 0.459 | 0.145 |
| | *(1.725)* | (1.230) | (0.146) | (0.137) | (0.876) | (0.588) | (0.281) | (0.202) |
| *lnNOLSShare$_{ijk}$* | *-1.751**** | -1.616*** | -0.154** | -0.248*** | -1.409*** | -0.635** | -0.190 | 0.0441 |
| | *(0.532)* | (0.399) | (0.073) | (0.082) | 0.470) | (0.309) | (0.193) | (0.166) |
| *lnOHSBShare$_{ij}$* | | 0.591** | | 0.481*** | | 0.402* | | 0.570*** |
| | | (0.268) | | (0.151) | | (0.241) | | (0.122) |
| *lnOMSBShare$_{ij}$* | | 0.080 | | -0.001 | | 0.014 | | -0.226 |
| | | (0.521) | | (0.334) | | (0.454) | | (0.301) |
| *lnOLSBShare$_{ij}$* | | -0.154 | | -0.163 | | -0.0230 | | -0.0226 |
| | | (0.272) | | (0.217) | | (0.271) | | (0.197) |
| *Observations* | 465 | 465 | 465 | 465 | 465 | 465 | 465 | 465 |

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.
N = 465 for all regressions. All regressions include the full list of covariates, countries and times fixed effect and a constant as shown in Table 3.6 and Table 3.7 for co-inventorship and co-authorship as dependent variables respectively. We add a unit to all of the above explanatory variables before logarithmic transformation in order to account of the presence of many zeros.

Another issue with our results concerns our focus of Indian and Chinese migrant, which was deliberately arbitrary and motivated only by the importance of these two diaspora groups. One may doubt of the interpretation we provide of our results, based as it is on a view of knowledge as tacit and of migrants' social networks as important knowledge carrier, to the extent that, if true and relevant, such results ought to hold also for other diaspora groups. Hence, we replicate our baseline exercise with ten other *NOHS* diasporas, all of which belong to the top *hs* migrants sending countries to the 31 OECD countries of our sample (see the list in Table A6 in Appendix A).

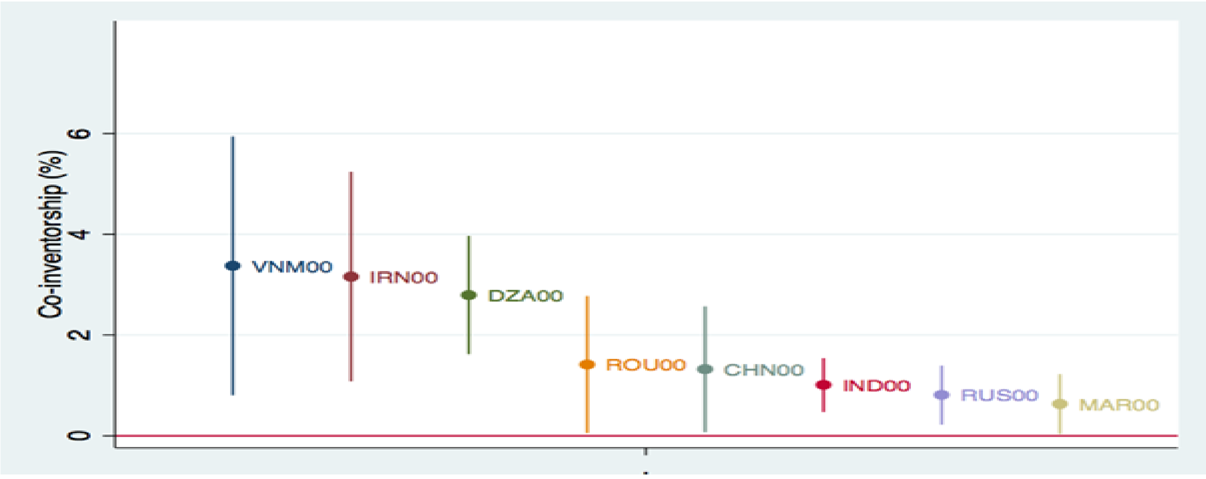
More specifically for each *NOHS* diaspora *k* among the ten, we compute the S&T collaborations elasticities with respect to the size of the product of *NOHS* migrant *k* shares in host countries *i* and *j* – $NOHSShare_{ijkt.}$ Following once more Felbermayr et al. (2010), we run separate regressions for each *NOHS* diaspora *k* and in each DIOC edition. Results from these regressions are summarized in Tables A4 and A5 Appendix A, for each S&T collaboration variable co-inventorship and co-authorship respectively. In what follows, we simply report estimated elasticities through a few graphs.

In general, we find that the positive results we got for the case of the Chinese and Indian *hs* diasporas extend to other *NOHS* diasporas, similar to what was found in the trade literature for overall migration (Felbermayr et al., 2010). Furthermore, for the diasporas shown in the below graphs, in most cases their effects are not significantly different one from another.

Figure 4 below reports the point estimates obtained for each *NOHS* diaspora variable from separate regressions – co-inventorship equation – as dots at the center of the spikes. The upper part represents DIOC 00/01 edition and the lower part DIOC 10/11. The spikes denote the 95% confidence intervals of each coefficient estimate – all shown estimates are statistically significant at least at the 1% level. The figure shows the Chinese and Indian *hs* migrants are not the most influential *NOHS* diasporas when it comes to co-inventorship. Indeed, the Vietnamese and Iranian *hs* diaspora effects on co-inventorship appear to be relatively higher in DIOC 00/01 and DIOC 10/11respectively –3.372 and 3.159. Iranian and Pakistanis *hs* migrants are found at the second position in DIOC 00/01 and DIOC 10/11 respectively, while Chinese and Indian *hs* migrants only win a fifth and sixth position respectively in DIOC 00/01. The marginal effects of these *NOHS* migrant variables are presented in Table A3 in Appendix A. This table shows Algerian *hs* to have the highest marginal effect on co-inventorship in DIOC 00/01 and DIOC 10/11 as in average an increase of their share in country-pair destination – by 0.007 and 0.008 respectively - would lead to an increase co-inventorship instances by 268 and 63 respectively.

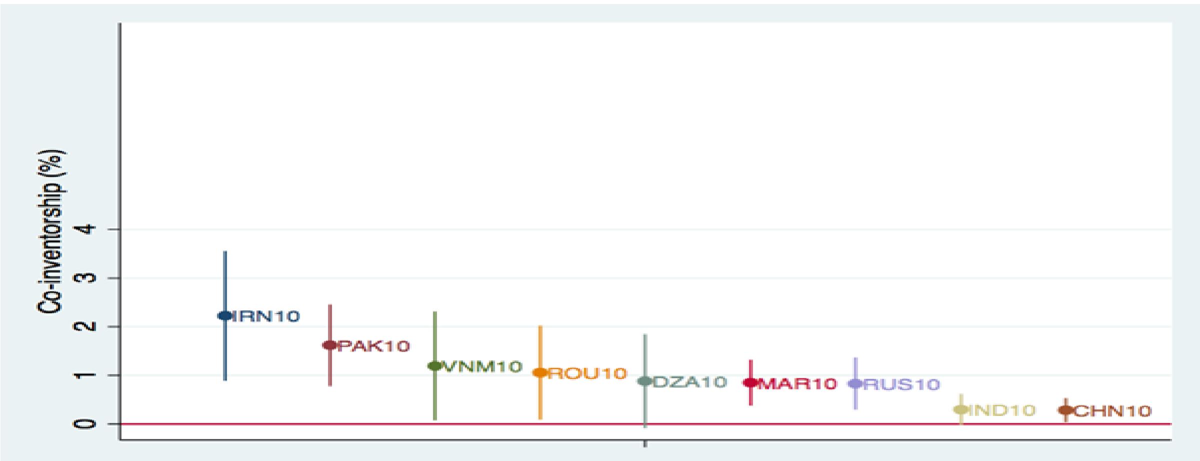
**Figure 4**: Co-inventorship elasticities of different *NOHS* diasporas

DIOC 00/01




VNM00 = Vietnamese *hs* diaspora DIOC 00/01 - IRN00 = Iranian *hs* diaspora DIOC 00/01 - DZA00 = Algerian *hs* diaspora DIOC 00/01 - ROU00 = Romanian *hs* diaspora DIOC 00/01 - CHN00 = Chinese *hs* diaspora DIOC 00/01 - IND00 = Indian *hs* diaspora DIOC 00/01 - RUS00 = Russian *hs* diaspora DIOC 00/01 - MAR00 = Moroccan *hs* diaspora DIOC 00/01 The dots represent the co-inventorship elasticities resulting from an increase in the *NOHS* diaspora *k* share by 1%.

**DIOC 10/11**

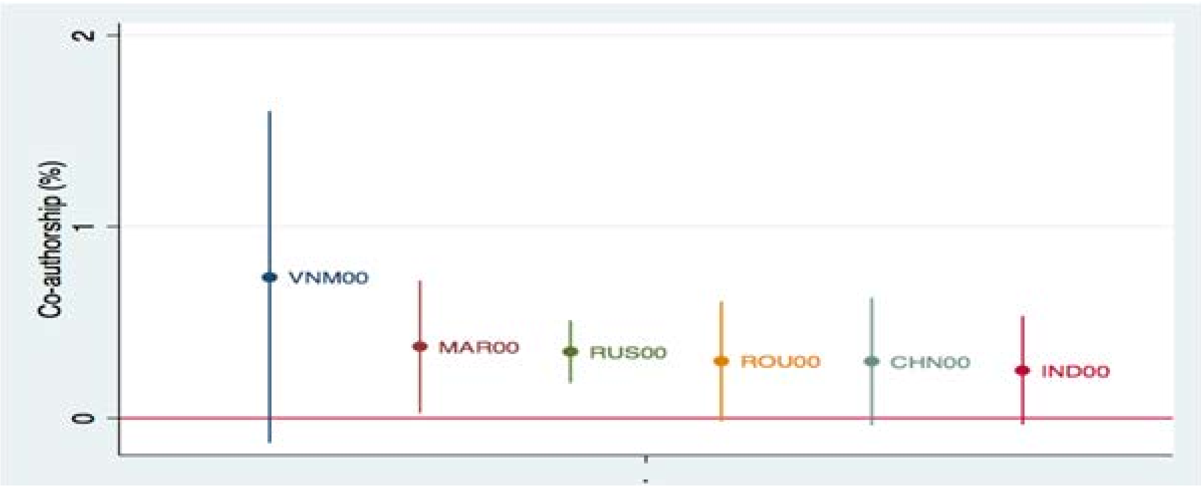



IRN10 = Iranian *hs* diaspora DIOC 10/11 -PAK10 = Pakistanis *hs* diaspora DIOC 10/11 -VNM10 = Vietnamese *hs* diaspora DIOC 10/11 -ROU10 = Romanian *hs* diaspora DIOC 10/11 - DZA10 = Algerian *hs* diaspora DIOC 10/11 - MAR10 = Moroccan *hs* diaspora DIOC 10/11 - RUS10 = Russian *hs* diaspora DIOC 10/11 -IND10 = Indian *hs* diaspora DIOC 10/11 -CHN10 = Chinese *hs* diaspora DIOC 10/11The dots represent the co-inventorship elasticities resulting from an increase in the *NOHS* diaspora *k* share by 1%.

The point estimates obtained for each *NOHS* diaspora from separate co-authorship regressions are shown in Figure 5 below. Similarly to co-inventorship regressions, all the reported estimates are statistically significant at least at the 1% level, with DIOC 00/01 edition in the top graph and DIOC 10/11 in the bottom one. Here again, the Chinese and Indian *hs* diasporas do not appear to have the biggest effect on co-authorship. Vietnamese *hs* diaspora induce the most important effect on co-authorship in DIOC 00/01 – 0.736 – while in DIOC 10/11 Pakistanis *hs* diaspora seem to entail the strongest effect– 0.653. Table A3 in Appendix A reports the marginal effect of Vietnamese *hs* diaspora on co-authorship in DIOC 00/01, which is, if their share in country-pair destination increases by 0.012, co-inventorship would increase by 300. While in DIOC 10/11, the marginal effect of Pakistanis *hs* diaspora denotes an increase of their share by 0.015 induces an increase in co-authorship by 797.

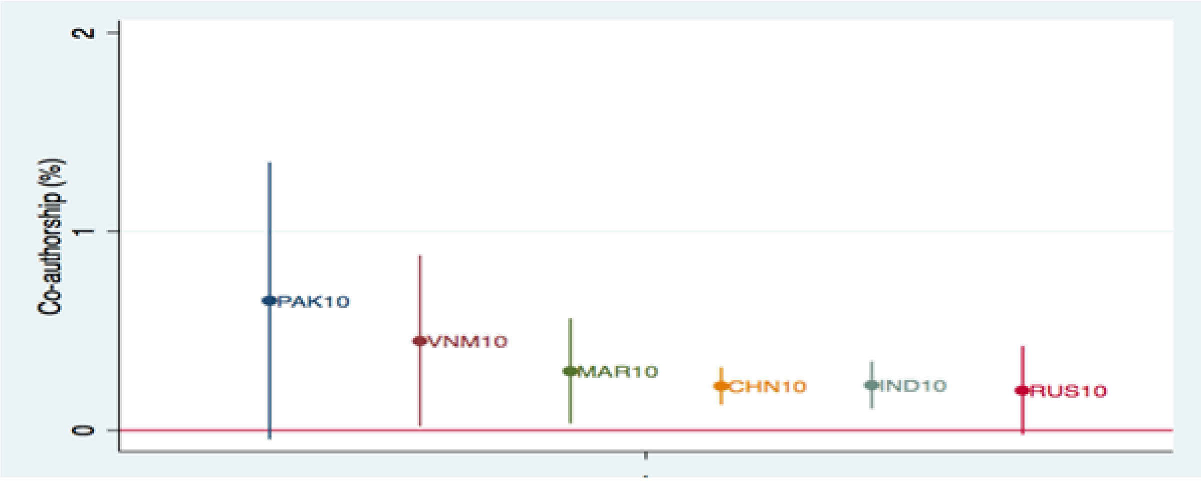
**Figure 5**: Co-authorship elasticities of different *NOHS* diasporas

DIOC 00/01




VNM00 = Vietnamese *hs* diaspora DIOC 00/01 - MAR00 = Moroccan *hs* diaspora DIOC 00/01 - RUS00 = Russian *hs* diaspora DIOC 00/01 -ROU00 = Romanian *hs* diaspora DIOC 00/01 -CHN00 = Chinese *hs* diaspora DIOC 00/01 -IND00 = Indian *hs* diaspora DIOC 00/01The dots represent the co-authorship elasticities resulting from an increase in the *NOHS* diaspora *k* share by 1%.

**DIOC 10/11**




PAK10 = Pakistanis *hs* diaspora DIOC 10/11 -VNM10 = Vietnamese *hs* diaspora DIOC 10/11 -MAR10 = Moroccan *hs* diaspora DIOC 10/11 -CHN10 = Chinese *hs* diaspora DIOC 10/11 -IND10 = Indian *hs* diaspora DIOC 10/11 -RUS10 = Russian *hs* diaspora DIOC 10/11.The dots represent the co-authorship elasticities resulting from an increase in the *NOHS* diaspora *k* share by 1%.

## 6 Concluding remarks

This paper aims at filling an existing gap in the *hs* migration literature by analysing the relationship between *hs* diasporas within specific receiving countries and S&T collaboration between these countries. More precisely, we have investigated the role played by two of the most influential non-OECD migrant groups – Chinese and Indian – on S&T collaboration across OECD destination countries. We have assessed two types of S&T collaborations: co-inventorship and co-authorship. Based on a gravity model, we have undertaken a cross-section analysis at a country-pair level, with our main explanatory variables capturing the potential size of social networks within each specific *hs* diaspora group. We find that both Chinese and Indian *hs* diasporas have a strong and positive impact on all our two dependent variables in both DIOC editions. Our results hold also when controlling for OECD *hs* migrant bilateral links, namely for links between the *hs* migrants from each pair of destination countries considered in our sample. In addition we find similar, and sometimes stronger effects after controlling for migration at two further skill levels, low and middle. This enables us to dismiss the possibility that our results might in fact capture some generic effects of migration in general, instead of some specific effects linked to *hs* migrants. In a further analysis, we replicate this exercise for 10 of the most important non-OECD *hs* diasporas besides the Chinese and Indian ones. Our findings suggest that although the latter are the most influential in absolute terms, there are other *hs* diasporas with positive and significant effects on S&T collaboration. In particular, we found comparable effects for the Iranian, Pakistani and Vietnamese *hs* diasporas.

Overall, our results hold after several robustness tests, suggesting a causality effect of the presence of *hs* diasporas on S&T collaboration among host countries. These results point to the importance of maintaining and strengthening linkages within *hs* diasporas abroad or internationally across destination countries. These linkages have the potential to favor knowledge exchange between destination countries through similar mechanisms as the ones suggested by the trade literature, such as lowering transaction costs and reducing informal barriers.

However, since our analysis was limited to a reduced list of *hs* diasporas within specific host countries we should refrain ourselves to generalize our results. Therefore, one interesting extension to this paper would be to conduct a similar analysis for the case of *hs* diasporas from developing countries within non-OECD countries destination countries – like for instance from Africa or developing countries in general. The proposed extension is relevant to the extent that South-South level of analysis is scant in this literature. Furthermore, it would be an important contribution to shift the attention from the '*brain drain'* narrative that has been at the core of the debate on South-North *hs* migration debate towards the '*brain gain'* potential of South-South *hs* migration

## References


Abramo, G., D'Angelo, C. A. and Solazzi, M. (2012), 'A bibliometric tool to assess the regional dimension of university-industry research collaborations', Scientometrics, Vol. 91, pp. 955-975.

Adepoju, A., T. Van Naerssen & A. Zoomers (2008), 'International Migration and National Development in SubSaharan Africa: View Points and Policies Initiatives in the Countries of Origin', in Aderanti et al. (Eds.). Brill E-Books. DOI:101163/ej.9789004163546.i-309.13.

Ackers, L. (2005),'Promoting Scientific Mobility and Balanced Growth in the European Research Area', Innovation 18, 301-17.

Agrawal, A., Cockburn, I. and McHale, J. (2006),'Gone but not forgotten: Knowledge flows, labor mobility and enduring social relationships', Journal of Economic Geography 6, 571-91.

Agrawal, A., Kapur, D., McHale, J.& Oettl, A., 2011. Brain Drain or Brain Bank? The Impact of Skilled Emigration on Poor-Country Innovation. J. Urban Econ. 69, 43–55. doi:10.1016/j.jue.2010.06.003

Appelt, S. et al. (2015), 'Which factors influence the international mobility of research scientists?, OECD Science, Technology and Industry Working Papers, 2015/02, OECD Publishing, Paris. http://dx.doi.org/10.1787/5js1tmrr2233-en

Bastian, D. (2006), 'Modes of knowledge migration: Regional assimilation of knowledge and the politics of bringing knowledge into the region'. *European Planning Studies*, *14*(5), 601-619.

Beaver, D. (2001). 'Reflections on scientific collaboration (and its study): past, present, and future', *Scientometrics*, *52*(3), 365-377.

Beine, M., S. Bertoli, & J. Fernandez-Huertas Moraga (2016) 'A practitioners' guide to gravity models of migration," The World Economy, Volume 39, Issue 4April 2016 .Pages 496–512.

Beine, M., Docquier, F.& Rapoport, H., 2001. Brain drain and economic growth: theory and evidence. J. Dev. Econ. 64, 275–289.

Belderbos, R., Carree, M. & Lokshin, B. (2004), 'Cooperative R&D and firm performance', Research Policy, 33, pp. 1477–1492.

Bergstrand, Jeffrey H. (1985), 'The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence', *The Review of Economics and Statistics* Vol. 67, No. 3 (Aug., 1985), pp. 474-481.

Breschi, S. & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of economic geography*, *9*(4), 439-468.

Breschi, S., F. Lissoni & E. Miguelez (2015), 'Foreign inventors in the US: Testing for Diaspora and Brain Gain Effects', paper presented at the 8th International Conference on Migration and Development, June 8-9, World Bank - Washington DC.

Breschi, S., F. Lissoni & G. Tarasconi (2014), 'Inventor Data for Research on Migration Innovation: A Survey and a Pilot', *WIPO Economic Research Working Paper,* 17, World Intellectual Property Organization –

Brown, M. (2002), 'Intellectual diaspora networks: their viability as a response to highly skilled migration', *Autrepart*, 22, 167–178.

Chellaraj, G., K. E.Maskus & A. Mattoo (2008), 'The Contribution of International Graduate Students to US Innovation', *Review of International Economics,* 16, 444-62.

Cincera, M., Kempen, L., Van Pottelsberghe, B., Veugelers, R., & Villegas, C. (2003), 'Productivity growth, R&D and the role of international collaborative agreements: some evidence for Belgian manufacturing companies'. *Brussels Economic Review*, *46*(3), 107-140.

Clarke, B. L. (1967, September), 'Communication patterns of biomedical scientists'. I. Multiple authorship and sponsorship of Federal Program volunteer papers. In *Federation Proceedings* (Vol. 26, No. 5, pp. 1288-1292).

Coffano, M. & G. Tarasconi (2014) 'CRIOS - Patstat Database: Sources, Contents and Access Rules', Technical Report 1, Center for Research on Innovation, Organization and Strategy.

Conchi, S. and Michels. C. (2014), "Scientific mobility – An analysis of Germany, Austria, France and Great Britain", Fraunhofer ISI Discussion Papers Innovation Systems and Policy Analysis No. 41, Karlsruhe, March 2014.http://www.isi.fraunhofer.de/isiwAssets/docs/p/de/diskpap_innosysteme_policyanalyse/discussionpaper_41_2014.pdf

De Backer, K., & Basri, E. (2008), 'The internationalisation of R&D'. *Staying Competitive in the Global Economy: Compendium of Studies on Global Value Chains*, 219-248.

Bresson, C. (1996), 'C. Economic interdependence and innovative activity', (Cheltenham: Edward Elgar)

Docquier, F., O. Lohest & A. Marfouk (2007) 'Brain Drain in Developing Countries'. ECON Working Papers; 2007/04 (2007)

Docquier, F., B. L. Lowell, and A Marfouk (2009)'A Gendered Assessment of Highly Skilled Emigration', *Population and Development Review* 35 (2): 297–321.

Docquier, F. & H. Rapoport (2012) 'Globalization, Brain Drain and Development', Journal of Economic Literature, 50, no. 3: 681-730.

Dosi, G. (1988),' Sources, procedures and microeconomic effects of innovation', Journal of Economic Literature,36, pp. 1126–1171.

Dunlevy, J. A. (2006). The influence of corruption and language on the protrade effect of immigrants: Evidence from the American States. *Review of Economics and Statistics*, *88*(1), 182-186.

Faems, D., Van Looy, B., & Debackere, K. (2005), 'Interorganizational collaboration and innovation: Toward a portfolio approach'. *Journal of product innovation management*, *22*(3), 238-250.

Felbermayr, G. J., Jung, B., & Toubal, F., (2010) 'Ethnic networks, information, and international trade: Revisiting the evidence'. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, 41-70.

Ferlbermayr, G. J. & F. Toubal (2012), 'Revisiting the Trade-Migration Nexus: Evidence from New OECD Data', World Development. Volume 40, Issue 5, May 2012, Pages 928–937.

Foley, C. F., & Kerr, W. R. (2013). Ethnic innovation and US multinational firm activity. *Management Science*, *59*(7), 1529-1544.

Gill, B. (2005),'Homeward bound? The experience of return mobility for Italian scientists', Innovation 18, 319-41.

Giovannetti, G., & Lanati, M. (2015) 'The indirect pro-trade effects of Indian ethnic networks', DEMO-India Research Report 2015/14.

Girma, S., & Yu, Z. (2000). The link between immigration and trade: evidence from the UK. Centre for the Research on Globalisation and Labour Markets (GLM). *GEP Research Papers*, (23).

Glänzel, W., & Schubert, A. (2004), 'Analysing scientific networks through co-authorship'. *Handbook of quantitative science and technology research*, *11*, 257-279.

Glänzel, W., Schubert, A., & Czerwon, H. J. (1999), 'A bibliometric analysis of international scientific cooperation of the European Union (1985–1995)'. *Scientometrics*, *45*(2), 185-202.

Gould, D. M. (1994). Immigrant links to the home country: empirical implications for US bilateral trade flows. *The Review of Economics and Statistics*, 302-316.

Guellec, D & van Pottelsberghe de la Potterie, B. (2001), 'The internationalisation of technology analysed with patent data', Research Policy 30 (2001) 1253–1266.

Granstrand, O., Håkanson, L., & Sjölander, S. (1993), 'Internationalization of R&D—a survey of some recent research'. *Research Policy*, *22*(5-6), 413-430.

Head, K., & Ries, J. (1998). Immigration and trade creation: econometric evidence from Canada. *Canadian journal of economics*, 47-62.

Herander, M. G., & Saavedra, L. A. (2005). Exports and the structure of immigrant-based networks: the role of geographic proximity. *Review of Economics and Statistics*, *87*(2), 323-335.

Hoekman, J., Frenken, K., & Tijssen, R. J. (2010), 'Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe', *Research Policy*, *39*(5), 662-673.

Hunt, J., & Gauthier-Loiselle, M. (2010). How much does immigration boost innovation?. *American Economic Journal: Macroeconomics*, *2*(2), 31-56.

Jaffe, A.B. (1986) 'Technological Opportunity and Flows of R&D: Evidence from Firms' Patents, Profits, and Market Value', Am. Econ. Rev. 76, 984–1001.

Jaffe, A.B., Caballero, R.J. (1993),'How high are the giants' shoulders: an empirical assessment of knowledge spillovers and creative destruction in a model of economic growth', in: Blanchard, Olivier, Fischer, Stanley (Eds.). MIT Press, pp. 15–86.

Javorcik, B.S.Ozden, C., Spatareanu, M. & Neagu, C., 2011. Migrant networks and foreign direct investment. J. Dev. Econ. 94, 231–24

Jones, B. F., Wuchty, S., & Uzzi, B. (2008), 'Multi-university research teams: Shifting impact, geography, and stratification in science'. *Science*, *322*(5905), 1259-1262.

Kamalski J. & A. Plume (2013), 'Comparative Benchmarking of European and US Research Collaboration and Researcher Mobility', A report prepared in collaboration between Science Europe and Elsevier's SciVal Analytics, September 2013. http://www.elsevier.com/__data/assets/pdf_file/0010/171793/Comparative-Benchmarking-o f European-and-US-Research-Collaboration-and-Researcher-Mobility_sept2013.pdf

Katz, J. S. (1994) 'Geographical proximity and scientific collaboration' *Scientometrics*, *31*(1), 31-43.

Kerr, W. (2009), 'The Agglomeration of US Ethnic Inventors' *NBER Working Papers.*

Kerr, W. R. (2008), 'Ethnic Scientific Communities and International Technology Diffusion', *Review of Economics and Statistics,* 90, 518-537.

Klein, P. and Ventura, G.J., 2007. TFP differences and the aggregate effects of labor mobility in the long run. *The BE Journal of Macroeconomics*, *7*(1).

Kugler, M. & Rapoport, H., 2007. International labor and capital flows: Complements or substitutes? Econ. Lett. 94, 155–162

Lata, R., T. Scherngell and T. Brenner (2012), 'Integration Processes in European R&D: A Comparative Spatial Interaction Approach Using Project Based R&D Networks and Co-patent Networks'

Laudel, G. (2003), 'Studying the brain drain: Can bibliometric methods help?', Scientometrics, Vol. 57 , pp. 215-237.

Leonard, D. & Sensiper, S. (1998), 'The role of tacit knowledge in group innovation', California Management Review, 40, pp. 112–132.

Luukkonen, T., Tijssen, R. J. W., Persson, O. and Siversten, G. (1993), 'The measurement of international scientific collaboration', Scientometrics, Vol. 28, pp. 15-36.

MacGarvie, M. (2006), 'Do Firms Learn from International Trade?', Rev. Econ. Stat. 88, 46–60.

Maggioni, M. A., M. Nosvelli & T. E. Uberti (2007) 'Space versus networks in the geography of innovation: A European analysis', Papers in Regional Science 86 (3), 471 – 493.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). Theories of international migration: A review and appraisal. *Population and development review*, 431-466.

Mayda, A. M. (2010) ‘International migration: a panel data analysis of the determinants of bilateral flows’, Journal of Population Economics, 23(4), 1249–1274.

Mayer, T. & S. Zignago. (2011) ‘Notes on CEPII’s Distances Measures: the GeoDist Database’. CEPII Work. Pap. 2011-25.

Meyer, J. P. & J. P. Wattiaux (2006), 'Diaspora Knowledge Networks: Vanishing Doubts, Increasing Evidence', in: International Journal on Multicultural Societies. *IJMS*, 8(1), 4-24.

Miguelez, E. (2016), 'Inventor Diasporas and the Internationalization of Technology', *The World Bank Economic Review*, lhw013.

Moed, H. & Plume, A. (2013), ‘Studying Scientific Migration in Scopus’, Scientometrics, Vol. 94, No. 3, pp. 929-942.

Moed, H. & Halevi, G. (2014), ‘A bibliometric approach to tracking international scientific migration’, Scientometrics, Vol. 101, pp. 1987-2001.

Moses, J.W. and Letnes, B., 2004. The economic costs to international labor restrictions: Revisiting the empirical discussion. *World Development*, *32*(10), pp.1609-1626.

Mountford, A. (1997). Can a brain drain be good for growth in the source economy?. *Journal of development economics*, *53*(2), 287-303.

Narin, F, Stevens, K. and Whitlow, E. S. (1991), ‘Scientific co-operation in Europe and the citation of multinationally authored papers’, Scientometrics, Vol. 21, pp. 313- 323.

Niebuhr, A. (2010). Migration and innovation: Does cultural diversity matter for regional R&D activity?. *Papers in Regional Science*, *89*(3), 563-585.

Nooteboom, B. (1999), ‘Innovation and inter-firm linkages: new implications for policy’. *Research policy*, *28*(8), 793-805.

Okubo, Y., & Zitt, M. (2004), ‘Searching for research integration across Europe: a closer look at international and inter-regional collaboration in France’. *Science and Public Policy*, *31*(3), 213-226.

Ortega, F. & G. Peri (2009) ‘The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005’, NBER Working Paper No. 14833. Issued in April 2009

Ortega, F., & G. Peri (2013) ‘The Role of Income and Immigration Policies in Attracting International Migrants’, Migration Studies, 1(1), 47–74.

Ottaviano, G. I. P. & G. Peri (2006), 'The economic value of cultural diversity: evidence fromUS cities', *Journal of Economic Geography,* 6, 9-44.

Özden, Ç., C. R. Parsons, M. Schiff, and T. L. Walmsley (2011) “Where on Earth Is Everybody? The Evolution of Global Bilateral Migration 1960–2000.” *The World Bank Economic Review* 25 (1): 12–56.

Ozgen, C., Nijkamp, P., & Poot, J. (2013). The impact of cultural diversity on innovation: evidence from Dutch firm-level data.

Pierson, A. S. and P. Cotegrave (2000), 'Citation figures suggest that the UK brain drain is a genuine problem', Nature, Vol. 407, p. 13.

Ponds, R., Van Oort, F., & Frenken, K. (2007). 'The geographical and institutional proximity of research collaboration', *Papers in regional science*, *86*(3), 423-443.

Rauch, J. E. (1999) 'Networks Versus Markets in International Trade', Journal of International Economics. Vol. 48. Issue 1. June 1999. Pages 7-35.

Rauch, J. E. & V. Trindade (2002), 'Ethnic Chinese Networks in International Trade', The Review of Economics and Statistics, 84, no. 1: 116-130.

Removille, J., & Clarysse, B. (1999), 'Intra-European scientific co-operation: measuring policy impact'. *Research Evaluation*, *8*(2), 99-109.

Santos-Silva, J.M.C., Tenreyro, S. (2006) 'The Log of Gravity', Rev. Econ. Stat. 88, 641–658.

Saxenian, A., Motoyama, Y., & Quan, X. (2002). Local and global networks of immigrant professionals in Silicon Valley. Public Policy Instit. of CA.

SCImago. (2007). SJR — SCImago Journal & Country Rank. Retrieved July 21, 2015, from http://www.scimagojr.com.

Squicciarini, M., H. Dernis and C. Criscuolo (2013), 'Measuring Patent Quality: Indicators of Technological and Economic Value', *OECD Science, Technology and Industry Working Papers*, No. 2013/03.

Stark, O., & Wang, Y. (2002). Inducing human capital formation: migration as a substitute for subsidies. *Journal of Public Economics*, *86*(1), 29-46.

Tinbergen, J. (1962), 'Shaping the World Economy: Suggestions for an International Economic Policy', New York: The Twentieth Century Fund.

Von Hippel, E. (1988), 'The Sources of Innovation', Nueva York: Oxford University Press.

Wagner, C. S. (2005), 'Six case studies of international collaboration in science', Scientometrics, Vol. 62, pp. 3-26.

Wagner, C. S. and Leydesdorff, L. (2005), 'Network structure, self-organization, and the growth of international collaboration in science', Research Policy, Vol. 34, pp. 1608-1618.

Widmaier, S. & J-C. Dumont (2011), 'Are recent immigrants different? A new profile of immigrants in the OECD based on DIOC 2005/06', OECD Social, Employment and Migration Working Papers No. 126, Directorate for Employment, Labour and Social Affairs, OECD Publishing http://dx.doi.org/10.1787/5kg3ml17nps4-en.

Windmeijer, F. A. G. & J. M. C. Santos Silva (1997), 'Endogeneity in Count Data Models: An Application to Demand for Health Care', *Journal of Applied Econometrics* 12 (3): 281–94.

Yoshikane, F. and Kagura, K. (2004), 'Comparative analysis of coauthorship networks of different domains: The growth and change of networks', Scientometrics, Vol. 60 , pp. 433-444.

## *Appendix A – Some tables*

**Table A1:** Detailed migration data sources by country of destination.

| DIOC 00/01 | | DIOC 10/11 | |
|---|---|---|---|
| **COUNTRY** | **SOURCE** | **COUNTRY** | **SOURCE** |
| Australia | Census, 2001 | Australia | Census, 2011 |
| Austria | Census, 2001 | Austria | European Labour Force Survey 2010/11 |
| Belgium | ESEG, 2001 | Belgium | Census, 2011 |
| Canada | Census, 2001 | Canada | National Household Survey (NHS) 2011 |
| Chile | Census, 2002 | Chile | The National Socio-Economic Survey, 2011 |
| Czech Republic | Census, 2001 | Czech Republic | Census, 2011 |
| Denmark | Register, 2002 | Denmark | Population Register 2011 |
| Estonia | Census, 2000 | Estonia | Census, 2011 |
| Finland | Register, 12/2000 | Finland | Population Register 2010 |
| France | Census, 1999 | France | Census, 2011 |
| Germany | LFS, 1998-2002, 2005 | Germany | Micro Census, 2011 |
| Greece | Census, 2001 | Greece | Census, 2011 |
| Hungary | Census, 2001 | Hungary | Census, 2011 |
| Ireland | Census, 2002 | Ireland | Census, 2011 |
| Israel | LFS, 2001 | Israel | Labour Force Survey 2011 |
| Italy | Census, 2001 | Italy | Census, 2011 |
| Japan | Census, 2000 | Japan | Census, 2010 |
| Luxembourg | Census, 2001 | Luxembourg | Census, 2011 |
| Mexico | Census, 2000 | Mexico | Census, 2010 |
| Netherlands | LFS, 1998-2002 | Netherlands | Census, 2011 |
| New Zealand | Census, 2001 | New Zealand | Census, 2013 |
| Norway | Registers, 12/2003 | Norway | Population Register 2011 |
| Poland | Census, 2001 | Poland | Census, 2011 |
| Portugal | Census, 2001 | Portugal | Census, 2011 |
| Slovakia | Census, 2001 | Slovakia | Census, 2011 |
| Slovenia | Census, 2002 | Slovenia | Census, 2011 |
| Spain | Census, 2000 | Spain | Census, 2011 |
| Sweden | Registers, 12/2003 | Sweden | Population Register 2010 |
| Switzerland | Census, 2000 | Switzerland | European Labour Force Survey 2010/11 |
| United Kingdom | Census, 2001 | United Kingdom | Census, 2011 |
| USA | Census, 2000 | USA | American Community Survey 2007-2011 |

*Notes: ESEG: Enquête socio-économique générale; LFS: Labour force survey.*

*Sources : DIOC 2010/11 methodology & DIOC 2000/01 methodology.*

**Table A2**: Descriptive statistics

| | DIOC 00/01 | | | | | DIOC 10/11 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | N | Mean | Std. Dev. | Min | Max | N | Mean | Std. Dev. | Min | Max |
| **OECD DESTINATION COUNTRIES** | | | | | | | | | | |
| Co-inventorship | 465 | 96 | 403.495 | 0 | 4928 | 465 | 72 | 279.75 | 0 | 3388 |
| Co-authorship | 465 | 407 | 980.492 | 0 | 9254 | 465 | 1220 | 2377.304 | 2 | 22334 |
| R&D cooperation | 465 | 212 | 505.732 | 0 | 5028 | 406 | 158 | 368.168 | 0 | 3055 |
| *Hs* bilateral shares | 465 | 0.092 | 0.653 | 0 | 9.684 | 465 | 0.089 | 0.527 | 0 | 7.294 |
| **CHINA** | | | | | | | | | | |
| *Hs* shares | 465 | 0.041 | 0.151 | 0 | 1.943 | 465 | 0.987 | 6.817 | 0 | 96.74 |
| **INDIA** | | | | | | | | | | |
| *Hs* shares | 465 | 0.071 | 0.265 | 0 | 2.414 | 465 | 0.331 | 1.198 | 0 | 12.046 |
| **RUSSIA** | | | | | | | | | | |
| *Hs* shares | 465 | 0.727 | 8.222 | 0 | 174.300 | 465 | 0.618 | 4.203 | 0 | 86.259 |
| **ROMANIA** | | | | | | | | | | |
| *Hs* shares | 465 | 0.078 | 0.372 | 0 | 6.484 | 465 | 0.126 | 0.322 | 0 | 3.548 |
| **VIETNAM** | | | | | | | | | | |
| *Hs* shares | 465 | 0.012 | 0.041 | 0 | 0.442 | 465 | 0.027 | 0.108 | 0 | 1.522 |
| **IRAN** | | | | | | | | | | |
| *Hs* shares | 465 | 0.021 | 0.049 | 0 | 0.429 | 465 | 0.034 | 0.069 | 0 | 0.665 |
| **MOROCCO** | | | | | | | | | | |
| *Hs* shares | 465 | 0.034 | 0.167 | 0 | 2.600 | 465 | 0.044 | 0.163 | 0 | 1.956 |
| **PAKISTAN** | | | | | | | | | | |
| *Hs* shares | 465 | 0.004 | 0.015 | 0 | 0.207 | 465 | 0.015 | 0.051 | 0 | 0.626 |
| **ALGERIA** | | | | | | | | | | |
| *Hs* shares | 465 | 0.007 | 0.041 | 0 | 0.445 | 465 | 0.008 | 0.039 | 0 | 0.478 |

**Table A3**: Marginal effects of co-inventorship and co-authorship

| | China | India | Russia | Romania | Iran | Morocco | Pakistan | Vietnam | Algeria |
|---|---|---|---|---|---|---|---|---|---|
| | **DIOC 00/01** | | | | | | | | |
| $\Delta$*Hs shares* | 0.041 | 0.071 | 0.727 | 0.078 | 0.021 | 0.034 | 0.004 | 0.012 | 0.007 |
| $\Delta$*Co-inv.* | 127 | 97 | 78 | 136 | 303 | 60 | ----- | 324 | 268 |
| $\Delta$*Co-aut.* | 119 | 100 | 147 | 121 | ----- | 152 | ----- | 300 | ----- |
| | **DIOC 10/11** | | | | | | | | |
| $\Delta$*Hs shares* | 0.987 | 0.331 | 0.618 | 0.126 | 0.034 | 0.044 | 0.015 | 0.027 | 0.008 |
| $\Delta$*Co-inv.* | 21 | 22 | 60 | 76 | 160 | 61 | 117 | 86 | 63 |
| $\Delta$*Co-aut.* | 270 | 279 | 245 | ----- | ----- | 365 | 797 | 550 | ----- |

**Table A4:** Other *NOHS* diaspora co-inventorship elasticities

| **Co-inventorship** | **DIOC 00/01 RUSSIA** | | | **DIOC 10/11 RUSSIA** | | | **DIOC 00/01 ROMANIA** | | | **DIOC 10/11 ROMANIA** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *lnNOHSShare*$_{ijk}$ | 0.808*** | 0.626*** | 0.422*** | 0.830*** | 0.530*** | 0.605*** | 1.414** | 1.122* | 1.012* | 1.057** | 0.883** | 0.993** |
| | (0.299) | (0.235) | (0.156) | (0.274) | (0.202) | (0.227) | (0.695) | (0.672) | (0.609) | (0.493) | (0.446) | (0.465) |
| *lnOHSBShare*$_{ij}$ | | 0.948*** | 0.569*** | | 0.776*** | 0.513** | | 0.799*** | 0.661*** | | 0.656*** | 0.562*** |
| | | (0.182) | (0.179) | | (0.139) | (0.224) | | (0.147) | (0.179) | | (0.138) | (0.214) |
| *Constant* | -59.51*** | -50.70*** | 9.127** | -2.960 | -2.775 | 7.908* | -4.334 | -8.539*** | 1.908 | -10.52*** | -34.86*** | -5.532 |
| | (11.42) | (2.684) | (3.624) | (2.294) | (4.166) | (4.239) | (3.626) | (3.061) | (4.869) | (2.873) | (3.349) | (3.861) |
| *Observations* | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 |
| *Pseudo R2* | 0.964 | 0.981 | 0.970 | 0.989 | 0.994 | 0.982 | 0.986 | 0.990 | 0.969 | 0.987 | 0.992 | 0.978 |
| | **DIOC 00/01 PHILIPPINES** | | | **DIOC 10/11 PHILIPPINES** | | | **DIOC 00/01 VIETNAM** | | | **DIOC 10/11 VIETNAM** | | |
| *lnNOHSShare*$_{ijk}$ | -0.343 | -0.413 | 0.138 | -0.172 | 0.274 | -0.037 | 3.372** | 3.428** | 2.533** | 1.192** | 0.530 | -0.274 |
| | (0.373) | (0.294) | (0.449) | (0.188) | (0.167) | (0.120) | (1.311) | (1.423) | (1.104) | (0.571) | (0.573) | (0.469) |
| *lnOHSBShare*$_{ij}$ | | 0.847*** | 0.796*** | | 0.649*** | 0.608** | | 1.172*** | 0.669*** | | 0.800*** | 0.676*** |
| | | (0.151) | (0.171) | | (0.155) | (0.208) | | (0.298) | (0.164) | | (0.149) | (0.210) |
| *Constant* | 20.26** | 1.075 | -1.036 | 8.126 | -64.67*** | 6.878 | -51.82** | -34.89 | 5.104 | -61.50*** | -67.14*** | -0.650 |
| | (8.664) | (6.362) | (5.509) | (25.70) | (24.73) | (9.152) | (22.44) | (26.04) | (4.572) | (19.35) | (20.13) | (7.072) |
| *Observations* | 461 | 461 | 431 | 461 | 461 | 431 | 461 | 461 | 431 | 461 | 461 | 431 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pseudo R2* | 0.987 | 0.991 | 0.967 | 0.986 | 0.991 | 0.976 | 0.934 | 0.948 | 0.969 | 0.987 | 0.993 | 0.973 |
| | **DIOC 00/01 IRAN** | | | **DIOC 10/11 IRAN** | | | **DIOC 00/01 MOROCCO** | | | **DIOC 10/11 MOROCCO** | | |
| *lnNOHSShare$_{ijk}$* | 3.159*** | 3.313*** | 2.568* | 2.224*** | 2.619*** | 1.749** | 0.628** | 0.466* | 0.519* | 0.851*** | 0.877*** | 0.786*** |
| | (1.062) | (0.955) | (1.356) | (0.679) | (0.866) | (0.763) | (0.303) | (0.273) | (0.282) | (0.240) | (0.223) | (0.231) |
| *lnOHSBShare$_{ij}$* | | 0.841*** | 0.707*** | | 0.566*** | 0.611*** | | 0.778*** | 0.668*** | | 0.373*** | 0.586*** |
| | | (0.152) | (0.168) | | (0.177) | (0.209) | | (0.141) | (0.164) | | (0.127) | (0.207) |
| *Constant* | -49.72*** | -51.73*** | 15.85*** | -76.97*** | -81.83*** | 12.95*** | -5.557 | -7.444** | -0.041 | -12.95*** | -24.69*** | -6.583*** |
| | (11.50) | (10.16) | (5.040) | (13.58) | (13.96) | (3.760) | (3.842) | (3.194) | (3.678) | (3.117) | (3.322) | (2.495) |
| *Observations* | 464 | 464 | 434 | 464 | 464 | 434 | 463 | 463 | 433 | 463 | 463 | 433 |
| *Pseudo R2* | 0.985 | 0.990 | 0.969 | 0.987 | 0.990 | 0.977 | 0.987 | 0.990 | 0.971 | 0.990 | 0.988 | 0.976 |
| | **DIOC 00/01 PAKISTAN** | | | **DIOC 10/11 PAKISTAN** | | | **DIOC 00/01 COLOMBIA** | | | **DIOC 10/11 COLOMBIA** | | |
| *lnNOHSShare$_{ijk}$* | -0.204 | -1.227 | -0.642 | 1.621*** | 0.896** | 1.041*** | -3.414 | -1.635 | -12.74 | 3.013 | -0.741 | -0.927 |
| | (1.436) | (1.575) | (1.295) | (0.429) | (0.384) | (0.383) | (4.748) | (4.524) | (9.329) | (3.003) | (2.459) | (3.295) |
| *lnOHSBShare$_{ij}$* | | 0.827*** | 0.697*** | | 0.730*** | 0.580*** | | 0.596*** | 0.828*** | | | 1.004*** |
| | | (0.148) | (0.166) | | (0.148) | (0.211) | | (0.190) | (0.286) | | 0.203 (0.274) | (0.359) |
| *Constant* | -30.31** | -33.98** | 10.52*** | -33.66* | -34.13* | 7.602* | -49.28*** | -48.73*** | 9.249** | -24.31* | -38.79** | -17.15 |
| | (15.37) | (14.27) | (3.753) | (20.31) | (20.09) | (4.151) | (11.24) | (11.07) | (4.594) | (13.50) | (18.95) | (11.72) |
| *Observations* | 464 | 464 | 434 | 464 | 464 | 434 | 465 | 465 | 436 | 465 | 465 | 436 |
| *Pseudo R2* | 0.986 | 0.990 | 0.968 | 0.988 | 0.993 | 0.977 | 0.987 | 0.988 | 0.991 | 0.972 | 0.927 | 0.919 |

| | DIOC 00/01 CUBA | | | DIOC 10/11 CUBA | | | DIOC 00/01 ALGERIA | | | DIOC 10/11 ALGERIA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lnNOHSShare$_{ijk}$* | -2.569 | -0.603 | -4.509 | -0.067 | 0.360 | 10.25 | 2.793*** | 2.435*** | 2.126*** | 0.881* | 0.459 | 0.347 |
| | (6.168) | (6.007) | (32.81) | (4.307) | (4.008) | (15.18) | (0.599) | (0.536) | (0.528) | (0.492) | (0.391) | (0.562) |
| *lnOHSBShare$_{ij}$* | | 0.620*** | 1.165*** | | 0.746*** | 0.773* | | 0.930*** | 0.725*** | | 0.822*** | 0.663*** |
| | | (0.176) | (0.313) | | (0.185) | (0.443) | | (0.166) | (0.166) | | (0.154) | (0.210) |
| *Constant* | -74.60*** | -72.92*** | -3.747 | -50.75 | -48.88 | -8.186 | -100.4*** | -99.67*** | 0.104 | -9.685** | -13.09** | -4.794 |
| | (25.40) | (25.35) | (6.389) | (51.26) | (48.68) | (8.360) | (13.13) | (11.06) | (4.298) | (4.561) | (5.642) | (3.305) |
| *Observations* | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 |
| *Pseudo R2* | 0.988 | 0.989 | 0.974 | 0.964 | 0.973 | 0.949 | 0.972 | 0.981 | 0.971 | 0.988 | 0.993 | 0.976 |

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

**Table A5:** Other *NOHS* diaspora co-authorship elasticities

| **Co-authorship** | **DIOC 00/01 RUSSIA** | | | **DIOC 10/11 RUSSIA** | | | **DIOC 00/01 ROMANIA** | | | **DIOC 10/11 ROMANIA** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *lnNOHSShare*$_{ijk}$ | 0.360*** | 0.282*** | 0.231*** | 0.201* | 0.166* | 0.192* | 0.297* | 0.240* | 0.362** | 0.136 | 0.0441 | 0.0602 |
| | (0.085) | (0.076) | (0.087) | (0.116) | (0.097) | (0.105) | (0.160) | (0.135) | (0.144) | (0.197) | (0.170) | (0.139) |
| *lnOHSBShare*$_{ij}$ | | 0.520*** | 0.383*** | | 0.449*** | 0.316*** | | 0.576*** | 0.466*** | | 0.473*** | 0.332*** |
| | | (0.080) | (0.099) | | (0.075) | (0.113) | | (0.098) | (0.096) | | (0.076) | (0.108) |
| *Constant* | -8.055*** | -7.842*** | -6.186*** | -7.531*** | -6.274** | -2.403 | -9.094*** | -9.317*** | -8.480*** | -9.690*** | -8.707*** | -3.357** |
| | (1.410) | (1.368) | (1.305) | (2.385) | (2.625) | (1.960) | (0.805) | (0.816) | (1.763) | (2.544) | (2.365) | (1.374) |
| *Observations* | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 |
| *Pseudo R2* | 0.975 | 0.981 | 0.977 | 0.973 | 0.982 | 0.980 | 0.972 | 0.977 | 0.976 | 0.972 | 0.982 | 0.979 |

| | **DIOC 00/01 PHILIPPINES** | | | **DIOC 10/11 PHILIPPINES** | | | **DIOC 00/01 VIETNAM** | | | **DIOC 10/11 VIETNAM** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *lnNOHSShare*$_{ijk}$ | -0.0318 | -0.0758 | 0.572*** | 0.0574 | -0.0176 | -0.00563 | 0.736* | 0.806** | 1.584*** | 0.451** | 0.324* | 0.174 |
| | (0.158) | (0.155) | (0.153) | (0.058) | (0.051) | (0.047) | (0.443) | (0.390) | (0.400) | (0.220) | (0.190) | (0.223) |
| *lnOHSBShare*$_{ij}$ | | 0.594*** | 0.475*** | | 0.477*** | 0.398*** | | 0.626*** | 0.507*** | | 0.445*** | 0.342*** |
| | | (0.092) | (0.092) | | (0.079) | (0.097) | | (0.089) | (0.093) | | (0.077) | (0.107) |
| *Constant* | -9.783*** | -10.30*** | -7.663*** | -10.34*** | -9.017*** | -5.312*** | -8.265*** | -6.859*** | -7.480*** | -9.768*** | -5.314** | -8.397*** |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (2.630) | (2.625) | (2.139) | (2.282) | (2.124) | (1.567) | (0.834) | (0.788) | (2.183) | (2.376) | (2.066) | (1.923) |
| *Observations* | 461 | 461 | 431 | 461 | 461 | 431 | 461 | 461 | 431 | 461 | 461 | 431 |
| *Pseudo R2* | 0.973 | 0.980 | 0.975 | 0.970 | 0.980 | 0.976 | 0.959 | 0.973 | 0.976 | 0.970 | 0.983 | 0.979 |
| | **DIOC 00/01 IRAN** | | | **DIOC 10/11 IRAN** | | | **DIOC 00/01 MOROCCO** | | | **DIOC 10/11 MOROCCO** | | |
| *lnNOHSShare*$_{ijk}$ | -0.124 | -0.323 | 0.238 | 0.464 | 0.433 | 0.821** | 0.373** | 0.285* | 0.306** | 0.299** | 0.261** | 0.325** |
| | (0.718) | (0.671) | (0.501) | (0.420) | (0.383) | (0.387) | (0.176) | (0.154) | (0.145) | (0.135) | (0.129) | (0.137) |
| *lnOHSBShare*$_{ij}$ | | 0.594*** | 0.489*** | | 0.471*** | 0.406*** | | 0.548*** | 0.419*** | | 0.366*** | 0.249** |
| | | (0.090) | (0.096) | | (0.073) | (0.097) | | (0.079) | (0.103) | | (0.084) | (0.115) |
| *Constant* | -9.929*** | -10.67*** | -7.437*** | -10.51*** | -8.974*** | -4.923*** | -9.556*** | -9.346*** | -6.438*** | -11.60*** | -11.52*** | -7.650*** |
| | (2.621) | (2.682) | (2.177) | (2.285) | (2.161) | (1.623) | (0.764) | (0.677) | (0.865) | (1.873) | (1.805) | (1.226) |
| *Observations* | 464 | 464 | 434 | 464 | 464 | 434 | 463 | 463 | 433 | 463 | 463 | 433 |
| *Pseudo R2* | 0.973 | 0.979 | 0.975 | 0.970 | 0.979 | 0.976 | 0.975 | 0.983 | 0.975 | 0.968 | 0.975 | 0.973 |
| | **DIOC 00/01 PAKISTAN** | | | **DIOC 10/11 PAKISTAN** | | | **DIOC 00/01 COLOMBIA** | | | **DIOC 10/11 COLOMBIA** | | |
| *lnNOHSShare*$_{ijk}$ | 0.356 | -0.892 | -0.844 | 0.653* | 0.160 | -0.0436 | -3.692 | -1.548 | 6.128 | -1.441 | -1.506 | -0.639 |
| | (0.994) | (1.204) | (0.520) | (0.356) | (0.244) | (0.210) | (3.519) | (3.302) | (3.886) | (1.134) | (0.924) | (0.789) |
| *lnOHSBShare*$_{ij}$ | | 0.589*** | 0.460*** | | 0.454*** | 0.329*** | | 0.485*** | 0.399*** | | 0.292*** | 0.183 |
| | | (0.075) | (0.091) | | (0.070) | (0.101) | | (0.082) | (0.102) | | (0.075) | (0.120) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Constant* | -7.136*** | -6.228*** | -5.010*** | -6.987*** | -5.889*** | -4.678*** | -13.44*** | -12.28*** | -10.42*** | -6.811*** | -13.76*** | -8.230*** |
| | (1.513) | (1.303) | (1.316) | (1.656) | (1.463) | (1.474) | (1.246) | (1.071) | (1.042) | (2.427) | (2.639) | (2.884) |
| *Observations* | 464 | 464 | 434 | 464 | 464 | 434 | 465 | 465 | 436 | 465 | 465 | 436 |
| *Pseudo R2* | 0.975 | 0.983 | 0.978 | 0.974 | 0.983 | 0.977 | 0.980 | 0.983 | 0.989 | 0.982 | 0.984 | 0.982 |
| | **DIOC 00/01 CUBA** | | | **DIOC 10/11 CUBA** | | | **DIOC 00/01 ALGERIA** | | | **DIOC 10/11 ALGERIA** | | |
| *$lnNOHSShare_{ijk}$* | -4.172 | -2.218 | 28.50*** | -1.496 | -0.878 | 14.63*** | 0.403 | 0.554 | 0.179 | -0.0754 | -0.0714 | 0.0555 |
| | (3.157) | (3.029) | (9.023) | (0.984) | (0.884) | (4.790) | (0.440) | (0.471) | (0.353) | (0.362) | (0.392) | (0.346) |
| *$lnOHSBShare_{ij}$* | | 0.439*** | 0.431*** | | 0.271*** | 0.220* | | 0.580*** | 0.482*** | | 0.324*** | 0.388*** |
| | | (0.095) | (0.107) | | (0.089) | (0.128) | | (0.079) | (0.101) | | (0.102) | (0.123) |
| *Constant* | -11.10*** | -10.59*** | -10.93*** | -11.57*** | -10.64*** | -12.02*** | -4.515** | -4.869** | -4.811** | -8.117*** | -7.551*** | -3.220** |
| | (1.690) | (1.741) | (1.332) | (2.502) | (2.414) | (2.637) | (2.254) | (2.112) | (2.253) | (1.936) | (1.738) | (1.409) |
| *Observations* | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 | 465 | 465 | 435 |
| *Pseudo R2* | 0.967 | 0.972 | 0.988 | 0.976 | 0.979 | 0.973 | 0.975 | 0.983 | 0.978 | 0.972 | 0.975 | 0.975 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

**Table A6**: Emigrant population 15+ in the OECD in 2010/11 by country origin

| COUNTRY OF ORIGIN | ISO 3 | Emigrant population ('000) | Highly educated emigrant population('000) |
|---|---|---|---|
| ALGERIA | DZA | 1,504 | 306 |
| **CHINA** | **CHN** | **3,862** | **1,655** |
| COLOMBIA | COL | 1,217 | 365 |
| CUBA | CUB | 1,205 | 345 |
| **INDIA** | **IND** | **3,441** | **2,080** |
| IRAN | IRN | 845 | 424 |
| MOROCCO | MAR | 2,630 | 392 |
| PAKISTAN | PAK | 1,088 | 378 |
| PHILIPPINES | PHL | 2,854 | 1,417 |
| ROMANIA | ROU | 2,643 | 483 |
| RUSSIA | RUS | 1,953 | 660 |
| VIETNAM | VNM | 1,879 | 524 |

*Source: DIOC 2010/11 http://www.oecd.org/els/mig/dioc.htm.*

## *Appendix B – IV strategy*

We examine the effect of Chinese and Indian *hs* migrants on co-inventorship and co-authorship after treatment for potential endogeneity through instrumental variable analysis. Drawing from (Miguelez, 2016), we use two instrumental variables: the size of the 1960 migrant group *k* in destination countries *i* and *j* (Özden et al., 2011) and the 1990 *ls* migrant group *k* in destination countries *i* and *j* (Docquier et al., 2009). We build the 1960 diaspora links for each specific migrant group *k* as the product of the 1960 migrant stocks within each destination country pair. The use of 1960 size is based upon two assumptions. On one hand, it is assumed that past immigration patterns to a specific destination strongly determine current immigration stocks and future flows in general. This in turn fuelled by network mechanisms, impacts on *hs* migration. While on the other hand, there should be no correlation between 1960 immigration stocks and current countries S&T collaboration as these immigration stocks record immigration flows from before the 1960s – a period of mass reconstruction within most OECD countries with a higher demand for less skilled labour. The second instrument is built as the product of the 1990 *ls* migrant group *k* stocks for each destination country pair. It represents a more contemporary variable given the period covered by our analysis. Again, here we assume past *ls* immigration to be determinant for current *hs* immigration stocks through networks operating within migrant communities, but also through training and education that *ls* migrants or their descendants might acquire at the destination country. On the contrary, there should be no correlation between past *ls* immigration and current countries S&T collaboration since in general there is no direct use of *ls* labour in S&T activities.

In Table B1 below, we show results from the GMM estimations of the PPML (Windmeijer & Santos Silva, 1997) for each dependent variable – co-inventorship and co-authorship – and for the Chinese and Indian *hs* migrants separately. A glance at the first column table values gives us an overview of the strength of our instruments. In columns (1) and (2), we see from the F-test statistics – which are 10.96 and 10.67 for DIOC editions 00/01 and 10/11 respectively – that the chosen instruments cannot be considered as weak. The Hansen J statistics for model specification test are reported at the bottom of columns (3) to (6). In all cases, our models seem not to exhibit any misspecification or over-identification issue. In general, the results suggest by omitting to account for endogeneity, we risk underestimating the real effect of *hs* migration on S&T collaboration as the estimates for each of the two *hs* diasporas considerably increase and remain significant as compared with results from the baseline models.

**Table B1**: GMM estimates with instrumented Chinese and Indian *hs* diasporas

| **CHINA** | First-stage | | GMM Co-inventorship | | GMM Co-authorship | |
|---|---|---|---|---|---|---|
| | **DIOC 00/01** | **DIOC 10/11** | **DIOC 00/01** | **DIOC 10/11** | **DIOC 00/01** | **DIOC 10/11** |
| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
| $lnNOHSShare_{ijk}$ | | | 2.489*** | 0.539*** | 0.402* | 0.313*** |
| | | | (0.796) | (0.191) | (0.212) | (0.102) |
| $lnNOSize60_{ijk}$ | 0.004*** | 0.029*** | | | | |
| | (0.001) | (0.008) | | | | |
| $lnNOLSSize90_{ijk}$ | 0.003*** | 0.024*** | | | | |
| | (0.001) | (0.008) | | | | |
| *Constant* | 0.317*** | 0.400** | -19.05 | -5.794 | -0.761 | -5.980*** |
| | (0.0495) | (0.203) | (14.86) | (4.691) | (1.116) | (1.208) |
| *F-test* | 10.14 | 20.18 | | | | |
| *p-value* | 0.000 | 0.000 | | | | |
| *Hansen's J chi2* | | | 1.608 | 10.062 | 9.191 | 9.760 |
| *p-value* | | | 0.658 | 0.122 | 0.163 | 0.135 |
| **INDIA** | | | | | | |
| $lnNOHSShare_{ijk}$ | | | 1.901*** | 0.653*** | 0.580*** | 0.449*** |
| | | | (0.389) | (0.140) | (0.130) | (0.096) |
| $\mathrm{lnNOSize60}_{\mathrm{ijk}}$ | 0.018*** | 0.032*** | | | | |
| | (0.003) | (0.007) | | | | |
| $lnNOLSSize90_{ijk}$ | 0.005* | 0.014** | | | | |
| | (0.003) | (0.006) | | | | |
| *Constant* | -0.274*** | -0.492*** | 6.537* | 3.077 | -8.950*** | -5.266*** |
| | (0.074) | (0.152) | (3.561) | (2.805) | (1.005) | (1.156) |
| *F-test* | 10.96 | 10.67 | | | | |
| *p-value* | 0.000 | 0.000 | | | | |
| *Hansen's J chi2* | | | 15.413 | 11.310 | 12.662 | 9.154 |
| *p-value* | | | 0.118 | 0.418 | 0.316 | 0.165 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.
We add a unit to all of the above explanatory variables before logarithmic transformation in order to account of the presence of many zeros. All regressions include the full list of covariates or controls, countries and times fixed effect and a constant. Total observations are N = 465.

## *Appendix C – Exploring R&D cooperation*

In an extension of our paper, we investigate each of our focal *NOHS* diaspora effect on one important S&T collaboration variable which is R&D collaboration. Unlike co-inventorship and co-authorship which entail formal output from S&T collaboration, R&D cooperation captures the initial collaborative process leading to S&T production. Therefore, R&D cooperation stands more as an input measure. This thus makes R&D cooperation an imperfect or remote proxy of collaboration. However, R&D cooperation networks are broader to the extent they reflect at the same time basic and applied knowledge (Lata et al., 2012). Although, there remains a risk of capturing joint R&D efforts that will fail to result to an innovation or a scientific production. Yet the knowledge externalities to the involved partners, that occur during the collaboration process is non negligible. More importantly, the R&D cooperation variable points to collaborations decided at an organizational level. This point marks one of the key difference between this variable and the two other dependent variables from a collaboration incentive point of view – since co-inventorship and co-authorship are rather done at an individual level. Therefore, performing this exercise would help us getting a better understanding of the mechanisms behind findings from the co-inventorship and co-authorship analysis. That is, it will help us drawing a comparison between company or institution-related effects – from R&D cooperation – against individual attached effects – co-inventorship and co-authorship.

Our data source for R&D cooperation is the EU Framework Programme for Research and Technological Development (FP). Between 1984 and 2013 there have been seven FP waves, but due to our explanatory variable data constraint, we only consider the last three ones (FP5, FP6 and FP7). The FP database reports all R&D alliances and joint ventures that have been made under the auspice of the FP at an international level. Entities taking part of such alliances are individuals but mostly firms from the private sector, universities and other public institutions from the European Union (EU) and also from the rest of the world. To compute the R&D cooperation dependent variable, we proceed similarly as we did for the co-inventorship dependent variables. We perform the absolute counting of the number of co-partners in a joint venture, per country pair, per year as illustrated in the tables C1 and C2 below for a sample of two projects registered under the FP5:

**Table C2**: Sample of R&D projects under the FP5

| Project id | Start Date | Contractor Country | FP |
|---|---|---|---|
| 51424 | 01/02/00 | DE; ES; ES; FR; IT; SE; SE | FP5 |
| 51426 | 01/02/00 | DE; IL; IT; SE | FP5 |

**Table C1:** Counting of country pairs R&D cooperation

| Ctry1 | Ctry2 | Year | R&D cooperation | FP |
|---|---|---|---|---|
| DE | ES | 2000 | 2 | 5 |
| DE | FR | 2000 | 1 | 5 |
| DE | IT | 2000 | 2 | 5 |
| DE | SE | 2000 | 3 | 5 |
| DE | IL | 2000 | 1 | 5 |
| ES | FR | 2000 | 2 | 5 |
| ES | IT | 2000 | 2 | 5 |
| ES | SE | 2000 | 4 | 5 |
| FR | IT | 2000 | 1 | 5 |
| FR | SE | 2000 | 2 | 5 |
| IT | SE | 2000 | 3 | 5 |
| IL | IT | 2000 | 1 | 5 |
| IL | SE | 2000 | 1 | 5 |

## C.1 Chinese and Indian *hs* links impacts on international R&D cooperation

In Table C3 below, we present the results from the same model specifications as with the co-inventorship and co-authorship dependent variables for explaining R&D cooperation with Chinese and Indian *hs* diasporas alternatively.

The first column illustrates the results from the baseline model in DIOC 00/01 with the Chinese *hs* diaspora variable as the main explanatory variable. We find a significant estimate of 1.245 for this variable. This result means if we double the sample probability of getting two Chinese *hs* from a random draw of two individuals from the total *hs* population of two host countries, the size of their

R&D cooperation would increase by of 124.5%. The marginal effects are computed using our sample average values for each of these variables. Therefore, an increase of the average country-pair Chinese *hs* share from 0.041 to a value of 0.082 would result to an increase in the R&D cooperation average value from 212 to the value of 476. Interestingly, we also find a similar positive effect of the estimate for the technological similarity variable – a value of 1.227. In sharp contrast, the estimate for the product of R&D capacity in two countries does not give conclusive results as it is not significant. As for the common gravity covariates, we don't find any effect of their estimates except for the distance which is negative and significant. The second column shows the results from the baseline model regression with one additional variable denoting the *OHS* migrant bilateral links. As shown in column (2), the estimate for the Chinese *hs* diaspora just moderately increases to 1.300 and remains strong, while the estimate for the *OHS* migrant bilateral links is positive but not significant. However, the US seems to account for a major part of the effect of the Chinese *hs* diaspora in DIOC 00/01 as the strength of this variable estimate drops significantly when removing all observations with that country, as shown in column (3). Results from DIOC 10/11 only slightly differ from DIOC 00/01 results as described earlier. Indeed, from the baseline model in column (4) we find a strong estimate of 0.444 of the coefficient for the Chinese *hs* diaspora variable. That is, doubling the probability of getting two Chinese *hs* from a random draw of two individuals from two host countrieswould induce a raise in the country pair R&D cooperation by 44.4%. The marginal effect of this explanatory variable is derived from its sample mean value and the R&D cooperation one which are 0.987 and 138 respectively. So if this Chinese *hs* share doubles to 1.974, there would be a raise of R&D cooperation to 199. Interestingly, when adding the *OHS* migrant bilateral links variable in the baseline regression we get a significant estimate of 0.176 of this variable's coefficient, while the estimate for our main dependent variable nearly remains unchanged – see column (5). In columns (6) we see the US as a destination country accounts for all the effect of the *OHS* bilateral links variable as its coefficient loses its significance.

In contrast with the results we have for the case China, our results fail to find any impact of Indian *hs* diaspora on R&D cooperation in both first and second DIOC editions.

**Table C3**: Chinese and Indian *hs* diaspora in international R&D cooperation

| ***R&D cooperation*** | DIOC 00/01 CHINA | | | DIOC 10/11 CHINA | | | DIOC 00/01 INDIA | | | DIOC 10/11 INDIA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *lnNOHSShare*$_{ijk}$ | 1.245*** | 1.300*** | 0.910* | 0.444*** | 0.422*** | 0.378*** | -0.231 | -0.254 | -0.508 | 0.040 | 0.043 | 0.074 |
| | (0.397) | (0.395) | (0.537) | (0.138) | (0.133) | (0.116) | (0.175) | (0.182) | (0.344) | (0.182) | (0.185) | (0.090) |
| *lnOHSBShare*$_{ij}$ | | 0.038 | -0.014 | | 0.176** | -0.008 | | 0.058 | -0.123 | | -0.109 | 0.149* |
| | | (0.062) | (0.080) | | (0.077) | (0.172) | | (0.060) | (0.079) | | (0.185) | (0.078) |
| *Common lang.* | 0.099 | 0.015 | 0.115 | 0.118 | -0.025 | 0.147 | 0.028 | 0.008 | 0.338*** | 0.015 | 0.037 | 0.013 |
| | (0.077) | (0.051) | (0.101) | (0.104) | (0.041) | (0.143) | (0.042) | (0.046) | (0.108) | (0.101) | (0.108) | (0.040) |
| *ln(distance)* | -0.209*** | -0.223*** | -0.204*** | -0.167*** | -0.183*** | -0.153*** | -0.225*** | -0.220*** | -0.156*** | -0.367*** | -0.369*** | -0.184*** |
| | (0.033) | (0.029) | (0.031) | (0.038) | (0.030) | (0.036) | (0.029) | (0.030) | (0.076) | (0.080) | (0.081) | (0.030) |
| *Contiguity* | 0.024 | 0.025 | 0.030 | 0.009 | 0.006 | 0.011 | 0.028 | 0.024 | -0.005 | -0.141 | -0.130 | -0.002 |
| | (0.034) | (0.032) | (0.035) | (0.040) | (0.030) | (0.040) | (0.032) | (0.032) | (0.090) | (0.105) | (0.105) | (0.032) |
| *Colony* | -0.059 | -0.033 | -0.075 | -0.106** | -0.079* | -0.067* | 0.003 | -0.014 | -0.259** | -0.297* | -0.302* | -0.048 |
| | (0.060) | (0.051) | (0.056) | (0.052) | (0.044) | (0.038) | (0.049) | (0.051) | (0.116) | (0.167) | (0.173) | (0.047) |
| *ln(RnD*$_i$**RnD*$_j$*)* | 0.027 | 0.018 | 0.005 | 0.083* | 0.097** | 0.024 | 0.016 | 0.015 | 0.198*** | 0.102*** | 0.100*** | 0.062*** |
| | (0.026) | (0.028) | (0.016) | (0.048) | (0.043) | (0.033) | (0.025) | (0.025) | (0.023) | (0.038) | (0.034) | (0.016) |
| *Tech. similarity* | 1.227*** | 0.780*** | 1.279*** | 0.772** | 0.278* | 0.754* | 0.877*** | 0.857*** | 7.286*** | 3.538*** | 3.600*** | 0.195 |
| | (0.458) | (0.243) | (0.482) | (0.346) | (0.149) | (0.449) | (0.250) | (0.247) | (0.730) | (0.531) | (0.541) | (0.156) |
| *Constant* | 2.125*** | 2.297*** | 2.058*** | 1.890*** | 1.855*** | 1.785*** | 4.302*** | 4.267*** | -0.206 | 4.219*** | 4.215*** | 4.537*** |
| | (0.401) | (0.333) | (0.385) | (0.345) | (0.302) | (0.341) | (0.397) | (0.398) | (0.881) | (0.976) | (0.976) | (0.415) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Observations* | 465 | 465 | 435 | 406 | 406 | 378 | 465 | 465 | 435 | 406 | 406 | 378 |
| *Pseudo R2* | 0.995 | 0.995 | 0.995 | 0.994 | 0.996 | 0.994 | 0.995 | 0.996 | 0.965 | 0.948 | 0.949 | 0.996 |
| *Countries & time FE* | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| *Without the US* | No | No | Yes | No | No | Yes | No | No | Yes | No | No | Yes |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

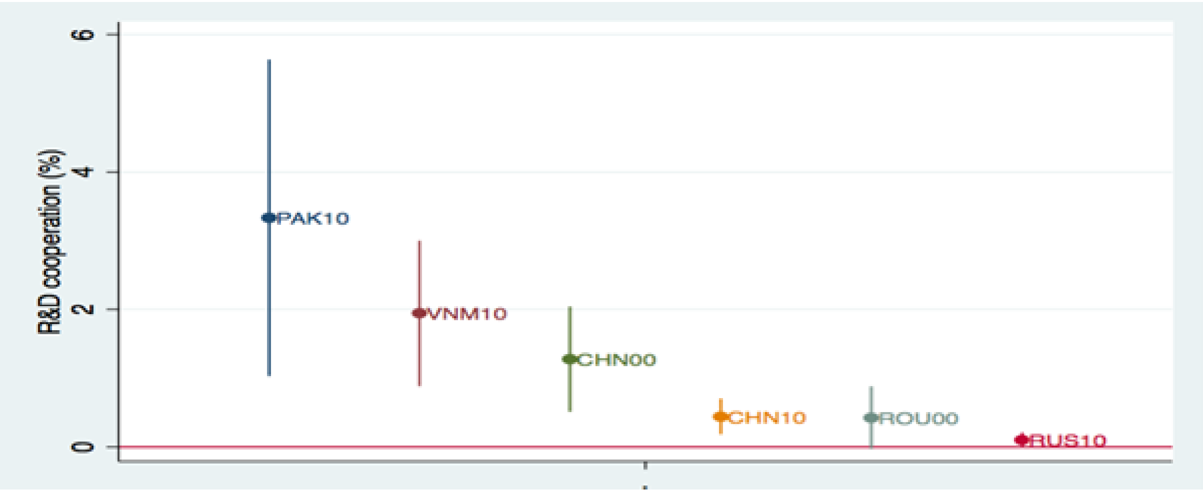
We add a unit to all of the above explanatory variables before logarithmic transformation in order to account of the presence of many zeros

## C.2 Other NOHS diasporas R&D cooperation elasticities

We run regressions for our baseline model for the top 10 most important *NOHS* diasporas within OECD. In Figure C1 below, we present the point estimates from separate R&D cooperation regressions obtained for each *NOHS* diaspora variables. Since we only find statistically significant estimates at least at the 1% level for two *NOHS* diasporas in DIOC 00/01 and four *NOHS* diasporas in DIOC 10/11- see Table C4 below –, we show the two DIOC editions in a single graph. Those point estimates represent elasticity effects. Pakistanis *hs* migrants are shown as the most influential *hs* diasporas in DIOC 10/11 in terms of their effect on R&D cooperation – 3.332 –, followed by Vietnamese *hs* diaspora at the second position.

**Figure C1**: R&D cooperation elasticities of different *NOHS* diasporas

DIOC 00/01 and DIOC 10/11




PAK10 = Pakistanis *hs* diaspora DIOC 10/11 -VNM10 = Vietnamese *hs* diaspora DIOC 10/11 - CHN00 = Chinese *hs* diaspora DIOC 00/01 -CHN10 = Chinese *hs* diaspora DIOC 10/11 -ROU00 = Romanian *hs* diaspora DIOC 00/01 -RUS10 = Russian *hs* diaspora DIOC 10/11 The dots represent the R&D cooperation elasticities resulting from an increase in the *NOHS* diaspora *k* share by 1%.

Overall, we get positive significant results for only few cases out of the 12 initial *NOHS* diasporas we have investigated. These results point to the nature of this dependent variable which is more representative of incentives or decisions to collaborate taken at an institutional level. Indeed, joint collaborations within the FP are likely to be initiated under an organizational setting and not by individuals. It is therefore unlikely that *hs* migrant exchanges which rather originate from *hs* migrants themselves, would have any significant effect on this variable.

**Table C4:** Other *NOHS* diaspora impact on R&D cooperation

| **R&D cooperation** | **DIOC 00/01 RUSSIA** | | | **DIOC 10/11 RUSSIA** | | | **DIOC 00/01 ROMANIA** | | | **DIOC 10/11 ROMANIA** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *$lnNOHSShare_{iik}$* | -0.270 | -0.196 | -0.0768 | 0.104** | 0.0764* | 0.0854* | 0.429* | 0.666*** | 0.849*** | 0.323 | 0.365 | 1.293*** |
| | (0.193) | (0.148) | (0.115) | (0.0477) | (0.0419) | (0.0494) | (0.233) | (0.238) | (0.252) | (0.283) | (0.302) | (0.271) |
| *$lnOHSBShare_{ij}$* | | -0.087 | -0.066 | | 0.146* | 0.139* | | -0.024 | -0.113 | | -0.142 | -0.004 |
| | | (0.063) | (0.093) | | (0.075) | (0.078) | | (0.095) | (0.081) | | (0.204) | (0.166) |
| *Constant* | 3.467*** | 5.149*** | 3.729*** | 4.738*** | 4.749*** | 5.203*** | 3.647*** | 3.676*** | 0.0537 | 3.761*** | 3.731*** | 2.506*** |
| | (0.884) | (0.797) | (0.575) | (0.363) | (0.362) | (0.383) | (0.779) | (0.798) | (0.871) | (0.863) | (0.861) | (0.946) |
| *Observations* | 465 | 465 | 435 | 406 | 406 | 378 | 465 | 465 | 435 | 406 | 406 | 378 |
| *Pseudo R2* | 0.987 | 0.992 | 0.993 | 0.996 | 0.996 | 0.996 | 0.957 | 0.963 | 0.973 | 0.947 | 0.948 | 0.959 |
| | **DIOC 00/01 PHILIPPINES** | | | **DIOC 10/11 PHILIPPINES** | | | **DIOC 00/01 VIETNAM** | | | **DIOC 10/11 VIETNAM** | | |
| *$lnNOHSShare_{iik}$* | -0.044 | -0.086 | -0.395 | 0.125 | 0.123 | 0.178 | -0.280 | -0.284 | -0.359 | 1.944*** | 1.924*** | 1.635*** |
| | (0.382) | (0.387) | (0.417) | (0.133) | (0.132) | (0.110) | (1.167) | (1.086) | (1.105) | (0.539) | (0.545) | (0.360) |
| *$lnOHSBShare_{ij}$* | | 0.043 | 0.062 | | 0.184** | 0.160** | | -0.097 | -0.058 | | -0.111 | 0.128 |
| | | (0.063) | (0.062) | | (0.080) | (0.080) | | (0.062) | (0.083) | | (0.113) | (0.080) |
| *Constant* | 2.302*** | 2.278*** | 3.994*** | 3.189*** | 3.129*** | 3.858*** | 3.961*** | 4.519*** | 1.246 | 3.180*** | 3.176*** | 5.941*** |
| | (0.486) | (0.487) | (0.352) | (0.454) | (0.457) | (0.349) | (0.497) | (0.652) | (0.967) | (0.476) | (0.474) | (0.388) |
| *Observations* | 461 | 461 | 431 | 402 | 402 | 374 | 461 | 461 | 431 | 402 | 402 | 374 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Pseudo R2* | 0.995 | 0.995 | 0.996 | 0.996 | 0.996 | 0.996 | 0.992 | 0.992 | 0.979 | 0.992 | 0.992 | 0.996 |
| | **DIOC 00/01 IRAN** | | | **DIOC 10/11 IRAN** | | | **DIOC 00/01 MOROCCO** | | | **DIOC 10/11 MOROCCO** | | |
| *$lnNOHSShare_{iik}$* | 0.669 | 0.667 | 0.798* | 0.428 | 0.422 | 0.720* | 0.123 | 0.121 | 0.128 | -0.036 | -0.041 | -0.011 |
| | (0.509) | (0.508) | (0.449) | (0.482) | (0.481) | (0.424) | (0.141) | (0.140) | (0.133) | (0.091) | (0.089) | (0.084) |
| *$lnOHSBShare_{ij}$* | | -0.056 | -0.062 | | 0.058 | 0.025 | | -0.050 | -0.058 | | 0.084 | 0.043 |
| | | (0.087) | (0.092) | | (0.132) | (0.148) | | (0.083) | (0.090) | | (0.128) | (0.150) |
| Constant | 4.364*** | 4.371*** | 3.351*** | 4.179*** | 4.185*** | 2.535*** | 4.723*** | 4.728*** | 3.587*** | 4.519*** | 4.532*** | 2.836*** |
| | (0.940) | (0.938) | (0.549) | (1.071) | (1.072) | (0.463) | (0.889) | (0.887) | (0.521) | (0.956) | (0.958) | (0.406) |
| Observations | 464 | 464 | 434 | 405 | 405 | 377 | 463 | 463 | 433 | 404 | 404 | 376 |
| Pseudo R2 | 0.992 | 0.992 | 0.993 | 0.994 | 0.994 | 0.994 | 0.992 | 0.992 | 0.992 | 0.994 | 0.994 | 0.994 |
| | **DIOC 00/01 PAKISTAN** | | | **DIOC 10/11 PAKISTAN** | | | **DIOC 00/01 COLOMBIA** | | | **DIOC 10/11 COLOMBIA** | | |
| *$lnNOHSShare_{iik}$* | 0.389 | 0.634 | 0.631 | 3.332*** | 3.234*** | 2.484** | -0.851 | -0.895 | 0.891 | -0.583 | -0.575 | 0.037 |
| | (1.083) | (1.122) | (1.163) | (1.173) | (1.125) | (1.094) | (1.415) | (1.417) | (1.413) | (0.373) | (0.376) | (0.753) |
| | | -0.036 | | | 0.180** | 0.163** | | 0.003 | 0.014 | | 0.023 | -0.246 |
| *Constant* | 4.237*** | 4.248*** | 6.145*** | 4.373*** | 4.369*** | 6.044*** | 2.704*** | 2.715*** | 2.309*** | 2.376*** | 2.369*** | 3.039*** |
| | (0.586) | (0.589) | (0.530) | (0.401) | (0.405) | (0.379) | (0.341) | (0.343) | (0.398) | (0.353) | (0.356) | (1.105) |
| *Observations* | 464 | 464 | 434 | 405 | 40 | 377 | 465 | 465 | 436 | 406 | 406 | 379 |
| *Pseudo R2* | 0.994 | 0.994 | 0.994 | 0.996 | 0.9 | 0.996 | 0.996 | 0.996 | 0.996 | 0.997 | 0.997 | 0.950 |

| | DIOC 00/01 CUBA | | | DIOC 10/11 CUBA | | | DIOC 00/01 ALGERIA | | | DIOC 10/11 ALGERIA | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *$lnNOHSShare_{iik}$* | -5.756 | -5.403 | -5.081 | -0.546 | -0.004 | 2.950 | 0.120 | 0.108 | 0.096 | 0.024 | -0.169 | -0.045 |
| | (4.099) | (4.086) | (6.855) | (2.263) | (2.008) | (2.443) | (0.374) | (0.376) | (0.381) | (0.320) | (0.415) | (0.326) |
| *$lnOHSBShare_{ii}$* | | 0.040 | 0.065 | | 0.212 | 0.200 | | -0.020 | -0.021 | | 0.081 | 0.165** |
| | | (0.129) | (0.133) | | (0.174) | (0.179) | | (0.078) | (0.081) | | (0.127) | (0.083) |
| *Constant* | 3.116*** | 3.115*** | 3.131*** | 2.005*** | 2.404*** | 2.401*** | 4.203*** | 4.210*** | 5.875*** | 3.765*** | 4.948*** | 5.231*** |
| | (0.934) | (0.936) | (1.070) | (0.673) | (0.421) | (0.509) | (0.584) | (0.589) | (0.531) | (0.439) | (0.976) | (0.380) |
| *Observations* | 465 | 465 | 435 | 406 | 406 | 380 | 465 | 465 | 435 | 406 | 406 | 378 |
| *Pseudo R2* | 0.972 | 0.972 | 0.972 | 0.958 | 0.982 | 0.982 | 0.994 | 0.994 | 0.993 | 0.995 | 0.994 | 0.996 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

# *Cahiers du GREThA*
# *Working papers of GREThA*


## GREThA UMR CNRS 5113

Université de Bordeaux

Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

**http://gretha.u-bordeaux.fr/**


## Cahiers du GREThA (derniers numéros – last issues)

*2018-04: DOYEN Luc, GAJARDOA Pedro : Viability standards and multi-criteria maximin*

*2018-05: EMLINGER Charlotte, LAMANI Viola: International Trade, Quality Sorting and Trade Costs: The Case of Cognac*

*2018-06: BONIN Hubert: Armateurs et négociants et la compétitivité de la cité-port de Bordeaux : les compagnies d'assurances maritimes dans les années 1830-1870*

*2018-07: V. TARIZZO, E. TROMEUR, O. THEBAUD, R. LITTLE, S. JENNINGS, L. DOYEN: Risk averse policies foster bio-economic sustainability in mixed fisheries*

*2018-08: ORSATTI Gianluca, STERZI Valerio: Do Patent Assertion Entities Harm Innovation? Evidence from Patent Transfers in Europe*

*2018-09: ROUGIER Eric, YOL Nicolas: The volatility effect of diaspora's location: A migration portfolio approach*

*2018-10: NAVARRO Noemí, H. Dan TRAN: Shock Diffusion in Large Regular Networks: The Role of Transitive Cycles*

*2018-11: CHIBA Fadoua, ROUILLON Sébastien: Intermittent electric generation technologies and smart meters: substitutes or complements*

*2018-12 : CHHORN Dina: Effect of Microfinance on Poverty and Welfare: New Evidence from 9 provinces in Cambodia*

*2018-13: BONIN Hubert: Les banquiers et les leçons de l'Histoire : lucidité ou aveuglement ?*

*2018-14: BALLET Jérôme : Anthropology and Economics: The Argument for a Microeconomic Anthropology*

*2018-15: SINCLAIR-DESGAGNE Bernard: The Preference for Monotone Decision Problems*


*La coordination scientifique des Cahiers du GREThA est assurée par Valerio STERZI.*
*La mise en page et la diffusion sont assurées par Julie VISSAGUET*