Inventor Migration and Knowledge Flows: a Two-Way Communication Channel?

Ernest Miguelez
GREThA CNRS UMR 5113, Université de Bordeaux
& AQR-IREA, University of Barcelona
ernest.miguelez@u-bordeaux.fr

Claudia Noumedem Temgoua
GREThA CNRS UMR 5113, Université de Bordeaux
ntclaud@yahoo.fr

Abstract
This paper documents the influence of networks of highly skilled migrants on the international diffusion of knowledge – particularly those with degrees and occupations in science, technology, engineering and mathematics. It investigates knowledge inflows to host countries brought in by skilled immigrants. It then explores knowledge feedback to home countries generated by these migrants. We test our hypotheses in a country-pair gravity model setting, for the period 1990-2010, using patent citations across countries to measure international knowledge diffusion. Our results confirm our hypotheses on the positive impact of skilled migrants on knowledge flows to host and home countries. However, they are not robust to instrumental variables and country-pair fixed-effects, and only matter in certain contexts: when the sending countries are developing nations and for knowledge diffusion within the boundaries of multinationals.

Keywords: STEM migration, knowledge diffusion, inventors, PCT patents, gravity models

JEL: C8, J61, O31, O33

1. Introduction

High-skilled workers are an important asset for a country's growth as they impact directly on knowledge production and diffusion (Nelson and Phelps, 1966; Vandenbussche et al., 2006). This is especially true of individuals with degrees and occupations in science, technology, engineering and mathematics (STEM), whose social and professional networks have been observed to disseminate important knowledge externalities (Moretti, 2004; Winters, 2014). Following the increasing globalization of STEM workers' mobility flows, special attention has been paid to the international dimension of such networks, with a focus on STEM migrants as key contributors to innovation in both their host countries (Chellaraj et al., 2008; Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010; Stephan and Levin, 2001) and in their countries of origin (Agrawal et al., 2011; Breschi et al., 2017; Kerr, 2008; Kuznetsov, 2006; Saxenian, 2006; Saxenian et al., 2002).

This paper aims to study the relationship between international knowledge diffusion and the migration of inventors, at a large, global scale. Migrant inventors constitute a representative category of STEM migrants - most of them R&D workers, highly involved in producing the knowledge that spurs economic growth and well-being. We test the hypothesis of a positive relationship between inventors' migration and knowledge flows in a gravity model framework for a sample of 33 OECD receiving countries and 133 developed and developing economies. We test whether the stock of migrant inventors originally from country i and resident in country j is positively associated with knowledge inflows (KI) into country j, originating from country i. We also test if the stock of migrant inventors originating from country i and resident in country j is positively associated with knowledge outflows (KO) to country i, originating from country j. The empirical analysis is made possible by the use of a novel dataset of inventors with information on both their residence and nationality (Miguelez and Fink, 2013). This information is available for a significant majority of Patent Cooperation Treaty (PCT) applications, from 1990 to 2010, thus making it unnecessary to estimate the probable ethnic origin of inventors.

A number of studies have already addressed similar issues. Most have focused on the US, however (Agrawal et al., 2011; Breschi et al., 2017; Ganguli, 2015; Kerr, 2008; Kerr and Lincoln, 2010; Moser et al., 2014). Systematic empirical evidence on the impact of migration on knowledge diffusion is still scarce. Such evidence is much needed as (1) it is well understood that the international diffusion of knowledge and new technologies is a source of economic growth and income convergence across countries (Eaton and Kortum, 1999; Keller, 2004), and (2) despite migrants' representing a small proportion of total worldwide population (around 3% - UN-DESA and OECD, 2013), the number of high-skilled, educated migrants (particularly STEM) in certain OECD countries has exploded in recent years (Kerr et al., 2016).

Baseline results point to a positive impact of high-skilled migrants on knowledge flows. We find that doubling the number of inventors of a given nationality in a destination country leads to a 5% increase in KI to that host economy. Equally, it produces a 5.4% increase in KO to their homelands.

We use different approaches to account for unobserved effects driving both talent and knowledge flows. First, we use a recently released index of migration policy to instrument our explanatory variable (Rayp et al., 2017). Second, we introduce country-pair fixed-effects (FE) to control for unobservables. Results of the effect of inventor migration on KO are consistent either with instrumental variables (IV) regressions or with country-pair FE estimates. Conversely, the effect on KI does not survive either of the two approaches.

We also test the existence of heterogeneous effects across broad technological fields, different groups of countries, and knowledge diffusion within multinational boundaries. We find that migrant inventors are important for KI originating in low- and middle-income countries only, as well as when diffusion occurs within organizational boundaries. Finally, we note that the effects diminish dramatically when the US and the BRICS countries are excluded from the analysis, proof of the importance of these countries as magnets for worldwide talent and as main providers of STEM migrants.

The rest of the paper is organized as follows: the next section summarizes the theoretical literature on highly skilled migration and innovation, and presents previous evidence on the topic. Section 3 focuses on the research

methods, including the description of our data and variables. Section 4 presents the results. We draw our conclusions in the last section.

2. Background

2.1. Theory and expectations

The international diffusion of ideas (especially from leading nations to poorer areas) is central to income convergence (Coe and Helpman, 1995; Eaton and Kortum, 1999; Keller, 2004). However, because what matters most from knowledge stocks is tacit in nature, it tends to resist diffusion (Audretsch and Feldman, 1996; Polanyi, 1958; Storper and Venables, 2004) and can only be transmitted by means of frequent face-to-face interactions and meetings. It thus requires "knowledge carriers" to transmit it over geographical distances (Breschi and Lissoni, 2009; Trippl, 2013). The international mobility of human capital has thus gained attention as a channel of international knowledge diffusion.

High-skilled migration can affect host country innovation through different channels. First, skilled immigrants directly contribute to the innovation activities of the receiving societies, simply because they add to the skilled labor force – quantitative contribution (Kerr, 2013). Second, as migrants tend to be positively self-selected, they specialize in jobs for which they have a comparative advantage with respect to native workers (Bosetti et al., 2015) – i.e., qualitative contribution (Kerr, 2017). Third, more migrants contribute to more culturally diverse societies, alongside the increased creativity and complexity that goes with it (Alesina et al., 2016; Bosetti et al., 2015; Ferrucci and Lissoni, 2019; Kemeny and Cooke, 2018). Fourth, they may also favor inward FDI (Hernandez, 2014) as well as cross-border acquisitions (Useche et al., 2019). This may affect the innovation potential of the firms involved. Finally, skilled immigrants are also sources of knowledge transfer by themselves, from their original countries to the host countries, as they bring new skills, abilities and ideas to the receiving society (Lissoni, 2018). They have the ability to transfer knowledge to their host country and to their firm that was previously locked within the cultural context of their homelands (Choudhury and Kim, 2019). These are precisely the ideas tested in the historical approaches we mention in the following section (Ganguli, 2015; Hornung, 2014; Moser et al., 2014), and the main idea we aim to test in the present paper. We therefore hypothesize that larger stocks of immigrants originating from country i who are residents in country j are likely to increase knowledge diffusion from origin country i to receiving country j.

Skilled migrants may not only contribute to innovation in their host country, but also to innovation in their homelands. A burgeoning body of literature has identified positive returns of migration for the countries of origin through diaspora networks. Diasporas have been defined as "part of a people (…) that maintains a feeling of transnational community among a people and its homeland" (Chander, 2001). This feeling can be exploited to the benefit of the home country. While most research has traditionally focused on monetary remittances, more recently, knowledge remittances have gained center stage too (Saxenian et al., 2002). Knowledge remittances may take two non-mutually exclusive forms: (1) skilled migrant workers maintaining personal and professional contacts with their home countries, favoring the diffusion of knowledge on a friendly or contractual basis (Breschi et al., 2017; Meyer, 2001; Meyer and Brown, 1999; Nanda and Khanna, 2010); (2) they may decide to move back to their home countries on a permanent or temporary basis, equipped with new skills and social networks (Baruffaldi and Landoni, 2012; Choudhury, 2016). We therefore expect that larger stocks of immigrants originating from country i who are residents in country j increase knowledge diffusion from host country j to origin country i.

Economic history has extensively documented skilled migration and subsequent knowledge diffusion, typically where the sending country has a technical advantage over the receiving country, at least in some fields – e.g. Germany with respect to the US in industrial chemistry in the 1930s (Moser et al., 2014). This is less the case for skilled migration nowadays, which increasingly comes from developing countries and is centered in English-speaking economies as hosts (Kerr et al., 2016). A large share of this migration is to complete graduate studies abroad, for instance (Breschi et al., 2018). We therefore expect migrant inventors to have a greater effect on KO than on KI.

Our empirical analysis also divides knowledge flow corridors between two groups, i.e., developed-developed countries vs developing-developed pairs. In principle, we expect the latter to affect KO particularly – developing countries benefitting most from having their diasporas abroad, while the former affect KI particularly – as South-North migration generally occurs in pairs where the sending country is not technically superior to the receiving country (with a large share of this migration occurring for study purposes). However, developed-developed country pairs are more technologically similar than developing-developed pairs. In such cases, developing-developed migrants could be more important for introducing novel ideas to their receiving societies.

In our analysis, we also differentiate between intra-company and inter-company knowledge diffusion. Indeed, the literature has long discussed the role of firms and multinational corporations (MNCs) in managing international knowledge transfer across different locations (Hedlund, 1986; Teece, 1977). While knowledge diffuses mainly locally (Audretsch and Feldman, 1996), the ability of MNCs to transfer knowledge more effectively than is possible through market-mediated channels, is a critical means of international knowledge diffusion (Hymer, 1976; Singh, 2008). This is nevertheless not easy, even within organizational boundaries, especially with regard to the cross-national transfer of complex or tacit knowledge (Kogut and Zander, 1993; Sorenson et al., 2006; Teece, 1977). The potential gains from accessing diverse knowledge hubs are often offset by difficulties in achieving integration of knowledge across multiple locations (Singh, 2008). In order to overcome the challenges in transferring knowledge across geographic distances, MNCs may rely on the mobility of their skilled employees between their countries of origin and their destination in the MNCs' location (Caligiuri and Bonache, 2016; Minbaeva and Michailova, 2004). As stressed by Kerr et al. (2016), the extent of employment mobility within the MNCs' boundaries is often ignored by the migration literature. Nonetheless, large MNCs may have almost half of their workforce employed outside the headquarters country – and likely to be moved around (possibly on a temporary basis), so the phenomenon is on the rise. Thus, we expect to find differences in the relation between knowledge flows and STEM migration within vs outside the firm's boundaries.

Finally, the diaspora literature has differentiated between direct and indirect effects of diaspora networks (Kapur and McHale, 2005). The former arise from diaspora members deliberately interacting with their home economies. The latter arise from diaspora members serving as intermediaries for easing knowledge transmission between migrants' home countries and third persons in their host economies. We expect direct effects of STEM migrants on KI and KO to be preponderant, but indirect effects are also likely to arise.

2.2. Previous evidence

In general, studies on high-skilled migration and innovation have long been confined to the area of economic history (Belfanti, 2006; Cipolla, 1972; Hornung, 2014; Luu, 2005). However, a group of scholars have worked on linking migration to innovation studies, mainly with the help of patent data (Agrawal et al., 2011; Breschi et al., 2017; Kerr, 2008; Kerr and Lincoln, 2010; Miguelez, 2018; Moser et al., 2014; Nathan, 2015). One stream of literature has focused on knowledge diffusion to receiving countries. This has generally been addressed by estimating the impact of the arrival of high-skilled workers on native knowledge creation: some papers have documented positive effects (Ganguli, 2015; Hunt and Gauthier-Loiselle, 2010), though some indicate small or even negative impacts (Borjas and Doran, 2015, 2012; Kerr and Lincoln, 2010). A critical challenge of this literature is identification, as most high-skilled workers may choose to relocate to highly-innovative, highly-rewarding places. Kerr and Lincoln (2010) apply a shift-share instrument across US States to study the impact of H1B visa admissions on local innovation. They find that skilled immigration leads to more patenting by inventors of Chinese and Indian origin, but not for natives – so immigrants contribute directly to innovation, rather than affecting native productivity through externalities. Using a historical migration exogenous shock, Moser et al. (2014) suggest that patenting by US-based inventors increased considerably in the 1930s in chemistry fields in which German Jewish émigrés were present, after being expelled from Nazi Germany. Interestingly, their results suggest that this effect was especially due to other inventors being attracted into the field, rather than an increase in the productivity of actual inventors. Ganguli (2015) is one of the few looking directly at migration and knowledge flows, rather than migration and innovation. The author exploits the fall of

the Soviet Union as a natural experiment and the sudden migration of Russian scientists to the US which resulted. She looks at a panel of US cities and scientific fields and shows a disproportionate number of citations to Soviet-era articles after the arrival of Russian migrants. Oettl and Agrawal (2008) identified internationally mobile inventors from the United States Patent and Trademark Office (USPTO) (when reporting different addresses in their patents) and find that the inventors' host countries gain knowledge inflows from their arrival, above and beyond the flows enjoyed by the firms recruiting them. Fassio et al. (2019) is one of the few studies, to our knowledge, that adopts an industry perspective – rather than a geographical one. The authors measure the impact of (skilled) immigration on innovation at industry level (citation-weighted patent production), which is critical, as skilled migrants tend to be concentrated in just a few industries. Indeed, in their analysis for France, Germany and the UK, the authors find heterogeneous effects across sectors, depending on their openness to trade and FDI.

From the perspective of migration-sending countries, the literature has usually depicted high-skilled migration as a source of brain drain, and hence political concern (Beine et al., 2001; Bhagwati and Hamada, 1974). More recently, studies claim that international co-ethnic ties may ease knowledge flows among high-skilled workers of the same origin back to the migrants' source country. Saxenian et al. (2002) surveys Silicon Valley scientists and engineers and discovers that around 82% of Chinese and Indians report having exchanged technical information with their peers back home, and 18% invest in their origin countries. Kerr (2008) uses patent data from the USPTO and by applying an ethnicity identification technique based on inventors' names shows that ethnic ties increase knowledge diffusion from the US to the migrants' home countries. The author estimates negative binomial models to show positive impacts of ethnic inventors in the US – seven foreign ethnicities identified – on knowledge flows back to their countries of origin, measured by patent citations. The effect is especially strong for high-tech industries and for the case of China, and the result is interpreted as evidence of positive returns for emigrants' sending countries. In a similar vein, Agrawal et al. (2011) build an Indian inventor database in USPTO patents using name identification techniques, and explore patent-level citations to study international knowledge flows from the US back to India. They find that patents by Indian inventors in the US do not seem to attract a higher-than-average rate of citations from the inventors' home country. The only (weak) exceptions are patents in Electronics, and patents owned by multinational firms. Interestingly, these results seem to suggest that the Indian diaspora is not a major source of knowledge feedback for the home country. The results of Agrawal et al. (2011) are reproduced and extended to another eight countries of origin in Breschi et al. (2017) – where migrant status is again identified by names. These authors find positive returns for emigrants' countries only in the case of China, Korea and Russia, and also for France, Italy and Japan within company boundaries – effects mediated by companies' self-citations. No results are found for Germany or India. They attribute the former to difficulties in determining who are true Germans residing in the US from their name and surname, a problem we do not share as we work directly with nationality. Finally, Oettl and Agrawal (2008) again provide evidence of a positive effect of internationally mobile inventors moving back to their source country. These authors find that the international movement of an inventor influences knowledge flows from the receiving country and the receiving firm to the source firm/country. And these effects double when the mover is hosted in a new geographic site within the same multinational company, in line with Breschi et al. (2017). This indicates that firms manage knowledge flows more effectively within their boundaries than outside them, and that mobility of labor reinforces intra-firm knowledge flows.

3. Methodology and data

3.1. Empirical approach

For the present analysis we use a standard gravity model – see Anderson (2011) for gravity models of trade and Beine et al. (2016) for gravity models of migration. Only a few studies have extended it to study knowledge diffusion patterns (Cappelli and Montobbio, 2016; Kerr, 2008; MacGarvie, 2005; Peri, 2005). The gravity model to be estimated for KI takes the following form:

$$KI_{ijt} = e^{\beta_0} . mig_{ijt-1}^{\beta_1} . Z_{ijt-1}^{\gamma_n} . e^{\tau_i} . e^{\tau_j} . e^{\delta_t} . \varepsilon_{ijt-1} \qquad ,$$

and for KO:

$$KO_{jit} = e^{\beta_0} . mig_{ijt-1}^{\beta_1} . Z_{ijt-1}^{\gamma_n} . e^{\tau_i} . e^{\tau_j} . e^{\delta_t} . \varepsilon_{ijt-1} \quad , \quad (2$$

where KI_{ijt} and KO_{jit} are, respectively, the total amount of knowledge flows from country i to country j in year t, and the number of knowledge flows from country j to country i, in year t. β_1 is our parameter of interest in both equations, mig_{ijt-1} is the number of active inventors of nationality *i* residing in country *j* during year *t*, Z_{ijt-1} is the set of dyadic and country-specific control variables in year *t*, and τ_i, τ_j and δ_t are country *i*, country *j* and time FE, respectively. ε_{ijt} stands as the error term. Note that all time-variant explanatory variables are lagged one year to minimize reverse causality problems.

When applying the gravity model we face the issue of strong skewness in the data distribution with relatively few high values at the bottom end. A common solution for this has been to transform the gravity equation into its logarithmic form – with a normal disturbance term, then to estimate it with OLS. However, this practice may result in some heteroskedasticity in the error terms, as pointed out by Santos Silva and Tenreyro (2006). Moreover, as our dependent variables contain a large number of zeros, their logarithmic transformation would be impossible without incurring serious bias due to arbitrary transformations (Burger et al., 2009). For these reasons, Santos Silva and Tenreyro (2006) recommend estimating the multiplicative form of the model using Poisson pseudo-maximum likelihood (PPML). Given all the above, we choose to apply the PPML regression to the conditional expectation of equations (1) and (2), as:

$$E(KI_{ijt}|X_{ijt-1}) = \exp[\beta_0 + \beta_1 \ln mig_{ijt-1} + \gamma_n \ln Z_{ijt-1} + \tau_i + \tau_j + \delta_t + \varepsilon_{ijt}] \quad (3$$

and

$$E(KO_{jit}|X_{ijt-1}) = \exp[\beta_0 + \beta_1 \ln mig_{ijt-1} + \gamma_n \ln Z_{ijt-1} + \tau_i + \tau_j + \delta_t + \varepsilon_{ijt}] \quad (4$$

3.2. Data

3.2.1. Patent citations and knowledge flows

Most studies reviewed in the previous sections use either trade flows or innovation outcomes after migration shocks as a proxy for knowledge exchange between countries. Finding a good measurement of the actual knowledge flows could be cumbersome to the extent that these flows are not tangible. The use of patent citations emerged as a way of overcoming this limitation – pioneered by Jaffe et al. (1993). Since then, this technique has been widely applied to various other studies, including migration research. Our solution, the use of citations as a proxy for knowledge flows, is not without its criticisms, most notably Jaffe and de Rassenfosse (2017) and Arora et al. (2018). However, these mostly relate to the use of citations as a proxy for inventor-to-inventor (or applicant-to-applicant) knowledge flows, while our analysis, at the aggregate, country-to-country level, aims to account for the outcome of a social, community phenomenon of migrant networking and communication. Researchers have suggested using applicant-added citations and disregarding examiner-added citations (Thompson, 2006), but this is not necessarily a good solution (applicant citations are actually added by attorneys). At the European Patent Office (EPO), for instance, the large majority of citations are added by examiners. Despite this, Duguet and MacGarvie (2005) find that EPO citations are good proxies for knowledge flows as measured by CIS data for a sample of French firms. The community idea makes it possible for inventor Z to receive a knowledge token from inventor A through a word-of-mouth process passing through inventors B, C, D and so forth. The origin of the flow may escape Z's attention (Z is unaware of A), but this does not mean that the transmission did not take place. Breschi and Lissoni (2005) develop this argument in full.

For their part, Thompson and Fox-Kean (2005) criticize the fact that citations might be a biased proxy of knowledge flows to the extent that they capture knowledge similarity rather than real knowledge flows. Indeed,

patents cite other patents within their technology far more frequently than those outside of their field. We address this issue, adding the appropriate controls and running separate regressions per broad technological field (see section 3.2.3).

Our dependent variable is built using cross-country citations to PCT patents – the patent database from the World International Patent Office (WIPO).[1] More precisely, we retrieve backward citations to PCT patents – as cited patents – from the OECD Citations database, July 2014, and geo-reference both cited and citing patents across all countries. From the initial data, only citing and cited patents with information on inventors and their countries of residence are selected, for the period 1990-2010,[2] and national-level citations are dropped. As some cited or citing patents are produced by teams of inventors scattered across 2 or more countries, citation counts are fractionalized as a function of the number of inventors in each citing and cited patent. Thus our dependent variable is just a dyadic variable returning the fractional count of backward patent citations from one country to another per year, weighted by the total number of inventors per country.

The KI dependent variable is built by counting the number of country j citations (citing patents) to country i patents (cited patents), grouped by country-pairs and year. The KO dependent variable is built by counting the number of country i citations (citing patents) to country j patents (cited patents), grouped by country-pairs and year. We remove self-citations at the inventor level from our analysis. Unfortunately, our sample of inventors is not disambiguated (it includes PCT inventors plus all cited inventors from any office). To exclude self-citations we compare the names of inventors listed in citing and cited patents, and exclude all citations with at least one inventor with the same (or similar) names.

We also divide knowledge flows between inter- and intra-firm citations. To identify intra-firm citations we incorporate information from the HAN OECD 2018 dataset which is a harmonized dataset of applicants (only name-harmonized, not disambiguated). This database incorporates information from ORBIS, which allows a benchmark name to link applicants' names. We also incorporate information from PATSTAT when the applicants of a given citing or a cited patent were not listed in HAN, as well as extensive manual checking. We then compare the names of citing and cited applicants, and identify as intra-firm those citations where citing and cited applicants have the same (or similar) names.

Next, we also separate citations from inventors whose country of origin is the cited one from the rest, in order to differentiate between direct knowledge flows (inventors in origin country citing only co-nationals in receiving country – KI, or migrants citing their home country's colleagues themselves – KO) and indirect knowledge flows (cross-nationality citations). This is by no means a straightforward task, as nationality is available for a large majority of PCT patents (citing ones) – but not all; and it is not available for the majority of cited patents which are not PCT. For a subsample of cited patents we were able to identify the nationality of the listed inventors if patents were either PCT or had a PCT as one of the members of the patent family. Whenever we cannot identify the full list of inventors' nationalities in a given citing-cited pair, we remove it completely from the analysis. For these reasons, results using this information should be treated with care.

3.2.2. STEM migration from inventor data

Most migration studies use education attained to determine skills level, and census data on the stocks of migrants with tertiary education as proxy for high-skilled migration. Yet, when it comes to STEM migration, data retrieved from censuses are less appropriate, as (1) education attained and skills can still differ markedly among

---

[1] The PCT is an international treaty administered by WIPO offering an advantageous route for seeking patent protection in more than one jurisdiction (in its contracting states). To seek for patent protection in multiple countries, applicants need to apply for patents in multiple offices. One simplifying route for doing this is offered by the PCT treaty, which applies a set of procedural rules common to all participating countries.

[2] Due to the relatively low quality and consistency of data prior to 1990 and after 2010, we focus on the period 1990-2010. The reason for the low quality of pre-1990 data is the limited use of the PCT system by worldwide applicants since the year of its inception (1978) until the early 1990s. Low consistency of post-2010 data is due to a lack of complete nationality information of inventors from around 2010-2012, as a result of the US enacting the Leahy-Smith America Invents Act, that resulted in the suppression of the requirement to list the nationality of the inventors in PCT documents (see Miguelez and Fink, 2013, for details).

tertiary educated workers, as this category collects people with science and engineering PhD together with people with non-STEM degrees or even non-university, tertiary education; (2) differences across countries emerge with respect to the quality of the education level attained, making cross-country comparisons troublesome; and (3) they are generally released every 10 years, which impedes longitudinal analysis in the short and medium run (and released to the public with a significant delay).

Some scholars have found a way to bypass these limitations by working with inventors as a proxy for STEM workers – a specific category of high-skilled migrants, most of them scientists and engineers. One advantage of using inventors' data in STEM migration studies is that migrant inventors stand as a more homogenous category of high-skilled migrants, highly involved in R&D, and behind the production of knowledge and new technologies.

For the present analysis we make use of a recent dataset on inventors from PCT patents (Miguelez and Fink, 2013), from which we are able to identify inventors with a migratory background on the basis of their nationality and place of residence. Compared to other inventor-based datasets, an advantage of using Miguelez and Fink (2013) is that it is the only one where information on inventors' nationality and residence are provided by inventors themselves.[3] It is therefore possible to identify migrant inventors by comparing information on nationality with that of residence. In our view, nationality is a more natural signal of origin than other proxies based on name identification techniques that have recently emerged in the literature (Agrawal et al., 2011; Breschi et al., 2017; Kerr, 2008). Additionally, patents administered by the PCT are international in nature, as applicants from all participant countries have, in principle, the same tendency to apply, contrary to other patent datasets which tend to be more biased towards one or other origin/destination country or region – the "home bias" effect.

Yearly country data on migrant inventors from the WIPO dataset are the starting point for computing our focal explanatory variable at a sending-receiving country pair level. More precisely, this variable stands as the annual number of active inventors – they appear listed in patents that year – who are nationals of an origin country i and residing in a given host country j (1990-2010).[4]

3.2.3. Control variables

Following related studies (Beine et al., 2016; MacGarvie, 2005; Miguelez, 2018; Peri, 2005), we control for geographical distance as well as cultural and historical ties between countries. Two variables are included for the former: (1) a dummy variable for contiguity, taking the value 1 if the two countries share a common border and 0 otherwise, and (2) a variable measuring the distance – in kilometers – between the capital cities of both countries. Cultural ties are proxied with a dummy for common language, taking the value 1 if both countries share at least one language and 0 otherwise. To control for historical ties, we include a dummy taking the value 1 if there has been a colonial link between the two countries and 0 otherwise.[5]

We also control for each country level of technological capacity with its total number of PCT inventors. This variable is informative to the extent that it measures the size of a country's innovation system, which determines both the amount of inflow and outflow of knowledge as well as the migration of talent in and out of the country. We also account for the fact that some countries are, on average, more cited than others, and may affect the direction of citation flows. We therefore introduce the host and home country average citation received per patent as controls.

---

[3]The PCT requires patent applicants to be a national of a PCT Contracting State, hence the requirement to file applicants' nationalities as well as their residences. And for international applications filed before 16 September 2012, inventors have to be listed as applicants for the purposes of the US designation. These two rules together made inventors' nationalities available for 80% of all PCT inventors.

[4] The migrant inventors variable is built by adding up patent-inventor pairs, per year, as PCT inventors are not disambiguated. If inventors acquire the host country nationality (and some of them apparently do), they are no longer counted as migrants. Manual checking indicates that this is a small phenomenon, so it is unlikely to affect our results to a large extent.

[5]All these control variables come from the *'Centre d'Etudes Prospectives et d'Informations Internationales'* (CEPII). (See Mayer and Zignago, 2011).

Additionally, other more economically-based country-pair variables are included to minimize bias in our focal coefficients due to confounding factors. First, we include an index of technological similarity between pairs of countries in order to control for whether they both share common fields of technological specialization. This index is computed using patent data from the EPO (Coffano and Tarasconi, 2014) and applying the following formula:

$$Tech.\,similarity_{ij} = \frac{\sum_{h=1}^{h=30} f_{ih} f_{jh}}{(\sum_{h=1}^{h=30} f_{ih}^2 \sum_{h=1}^{h=30} f_{jh}^2)^{1/2}}, \tag{5}$$

where f_{ih} stands for the share of patents of technological class h – according to the 30-class reclassification of IPC codes - held by country i, and f_{jh} the share of patents of technological class h held by country j.[6] Values of the index close to one indicate that a given pair of countries are technologically similar, and values close to zero indicate they are technologically remote from each other (Jaffe, 1986). With the inclusion of this control, we aim to tackle Thompson and Fox-Kean's (2005) criticisms on the tendency for there to be more citations within than across technological fields.

Second, we use trade flows (exports and imports), collected from the COMTRADE database, to proxy for economic integration between pairs of countries, as well as to account for knowledge diffusion embodied in goods and services (Bahar and Rapoport, 2018).

Finally, we include the stock of college-educated migrants from country i living in country j, taken from the 2000 census (Artuç et al., 2015). Descriptive statistics and the correlation matrix are presented in Appendix A.2.

Many of our explanatory variables contain zeros, and therefore their logarithmic transformation is problematic. To remedy this, we apply the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990). It behaves like a log-transformation, but is defined at zero (see recent applications by Bahar and Rapoport, 2018). Except for very small values of the variable, the inverse sine is approximately equal to its logarithmic version, and therefore it can be interpreted similarly. We apply this to all our explanatory variables except dummies, the index of technological similarity (ranging 0-1) and distance between capitals, which is log transformed.

4. Results

4.1. Stylized facts

Our regressions include pairs of countries formed from 33 OECD destinations, and 133 developed and developing sending economies – see the list of countries included in Appendix A.3.

Table 1 shows the top 20 largest players when it comes to knowledge flows and inventor migration, respectively – selected time period 2006-2010. From Table 1 (top panel) we can see that the largest flows of citations are amongst technology-leading or high-income countries, with the US being the largest origin of knowledge compared with other high-income nations, and China. Unsurprisingly, there is a strong reciprocity in the flow of knowledge to the extent that many top knowledge-recipient countries are also a source of knowledge for the countries they get knowledge from. Note that the largest flows of knowledge are exchanged amongst a small group of countries (13), most of them being high-income except for China. When dropping high-income countries from the list of knowledge-recipient countries, China appears as an important knowledge destination,

---

[6]This 30-class re-classification of IPC codes was originally proposed by the OST (Observatoire des Sciences et Techniques). For more details see Coffano and Tarasconi (2014).

receiving from the US alone around 20% of the share of citations going to low- and middle-income countries (Table A.2.1 in Appendix). This table also shows India as a second leading knowledge-recipient country.

[Table 1 about here]

The second panel splits citations between inter- and intra-company flows. At first sight, no major differences emerge between the two rankings (aside from the imbalance in the number of absolute citations between inter- and intra-company citations, as expected). Only the China-US corridor enters the top10 when intra-company citations are considered, though it already departed from the 11th position.

The bottom part of Table 1 looks at inventor migration corridors, which coincide to some extent with citation corridors, although with slight but striking differences. Unsurprisingly, the US appears as the most common host country for migrant inventors from 14 origins. Migrant inventors from China and India to the US account for 24% or almost one fourth of all migrant inventors in our dataset. There are also a large number of migrant inventors from Europe residing in the US, mainly from the UK, Germany and France, all of them are technology leading countries. When we focus on low- and middle-income sending countries, there is more variety in migrant inventors' origins (see Table A.2.2 in Appendix), but with the US as the main host economy. Migrant inventors coming from China and India to the US account altogether for around 57% of all migrant inventors originating from low- and middle-income countries.

4.2. Econometric results

Table 2 shows baseline regressions with the usual gravity variables as controls, plus the number of inventors in countries i and j to account for size and innovativeness (results using number of patents instead are qualitatively the same), in columns 1 and 2 for, respectively, KI and KO. Regressions also include country i FE, country j FE, and time FE. Results for control variables are for the most part significant and with the expected sign, with the exception of contiguity and same colonial past, which are not significantly different from zero.

The focal variable – migrant inventors – is positive and significant in the case of KI and KO. Doubling the number of inventors of a given nationality to a destination country leads to an 8.4% increase in KI to the host economy, while a similar increase in the number of migrant inventors increases KO by 8.7% - coefficients can be read as elasticities (Santos Silva and Tenreyro, 2006).

Columns 3 and 4 mimic 1 and 2 but add important country-proximity controls that are not accounted for in the usual gravity models, such as technological proximity, trade (exports and imports), and the stock of college-educated migrants originating from country i living in country j (2000 round census). The variables are positive and significant, and our focal variable diminishes its point estimate (though remains strongly significant in both cases) confirming the necessity of adding these three controls – 5% for KI, 5.4% for KO.

[Table 2 about here]

4.3. Confounding factors

Endogeneity issues could affect our baseline regressions and bias the results. Focal coefficients in KI regressions could be upward biased if more innovative (and highly-cited) receiving countries were to attract more inventors from abroad – KI and talent inflows would be spuriously correlated. For coefficients in KO regressions, they could be biased upwards if human capital and technological developments of sending countries increase their knowledge attractiveness (and citations received), and simultaneously increase the number of outward skilled migrants and the brain drain (see Clemens, 2014, for a discussion on the unexpected effects of development on the brain drain). Conversely, if unobserved technological development of sending countries reduces the emigration of STEM workers, our baseline estimates would be downward biased.

We address this using instrumental variables regressions. In particular, we use an index of migration policy as instrument, taken from Rayp et al. (2017). The authors compute a quantitative indicator of migration policy that accounts for restrictiveness of entry policy, staying requirements and regulations to foster integration. They combine publicly available data sources to provide a measure of "openness" to migrants (the larger the index, the more open the countries are to migration) based on these three concepts, for 38 countries between 1996 and 2014.

Note that, given that our sample ranges from 1990 to 2010, we do not use the years 2011-2014. In addition, the index is introduced with a 5-year time lag with respect to the dependent variable, and therefore a 4-year lag with respect to the variable to be instrumented (migrant inventors). The time needed for migration policy to affect inventor migration and their subsequent inventions (as we only observe migrant inventors when they patent, which could be some years after their arrival) is not immediate, and therefore a time lag is justified. We run IV regressions with different time lags of the instrument and we choose the time lag with the largest F-stat in the first stage. Consequently, our sample is reduced to the years 2001-2010 only.

As can be seen in Table 3, the instrument in the first stage (column 1) is positive and significant, as expected. Also, from the bottom of column 1 we learn that the F-stat of the first stage is well above 10, which is also a good sign of the appropriateness of the instrument. In columns 2 and 3 of Table 3 we reproduce our baseline regressions shown in Table 2, but for the indicated sample only, for comparison purposes. As can be seen, positive and significant coefficients remain. For presentation purposes, all the tables from now on do not show the coefficients for control variables, although these are always included (they are listed in the table notes). Results showing all controls can be requested from the authors.

Columns 4 and 5 show the result for KI and KO respectively. From column 4 we learn that, when our focal variable is instrumented, inventor migration does not influence KI to the host countries any more, indicating that baseline regressions were upward biased. In column 5, on the contrary, we see that the IV coefficient for KO increases considerably, and continues to be strongly significant, indicating that baseline regressions were downward biased.

[Table 3 about here]

In order to avoid losing too many observations, we take an alternative approach to deal with unobservables – in this case, country-pair unobserved factors.[7] Thus, in Table 4 we introduce country-pair FE and repeat our baseline regressions. Of course, due to the inclusion of pair FE we remove all time invariant variables. Columns 1 and 2 show the results for KI and KO respectively. Overall, results confirm our previous IV approach: positive and significant effects of inventor migration on KO, and inexistent on KI. This leads us to conclude that inventor migration favours the overall transfer of knowledge back to their homelands, but it does not seem to affect knowledge flows into the receiving countries.[8]

[Table 4 about here]

Table 5 goes one step further and shows pair-wise regressions adding interactions with our focal variable, migrant inventors, and 5 of our controls, namely, Technological similarity, Contiguity, Colony, Language, and Geographical distance (which we turn into the inverse of distance, proximity, for interpretation purposes). In their role of facilitating knowledge diffusion across borders, we may expect their impact to be larger for country pairs exhibiting stronger informational frictions, that is when the cognitive, cultural, or geographical distances between the two are more acute. Negative and significant interaction coefficients could be interpreted as causal evidence between inventor migration and knowledge diffusion, because if confounding effects drive both phenomena, they should work in such a way that they are not only capable of explaining the direct migration-

---

[7] The number of observations in the country-pair FE regressions decreases with respect to Table 2 due to the fact that non-linear models (e.g., Poisson) remove country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included.

[8] The country-pair FE approach is adopted in the rest of the paper, as we lose fewer observations than with the IV approach.

diffusion link, but also their different effects across several country-pair dimensions (Kugler et al., 2018; Miguelez, 2018). Results partially go in this direction, as most of the coefficients show negative signs, though they are only significant for technological and geographical proximities.

[Table 5 about here]

In sum, going back to section 2, it seems that results for KI do not align with Ganguli's (2015) findings on Russian scientists fleeing to the US, though her analysis of this highly specific context makes comparisons difficult – historical shock, scientists. Similarly, they do not coincide with Moser et al. (2014). Interestingly, however, these authors suggest that increases in innovation come from crowding-in effects, and are not due to native inventors' productivity shifts. Similarly, Kerr and Lincoln (2010) find positive effects of H1B visa holders' inflows on patenting, but all attributable to Chinese and Indian ethnic inventors, and not to natives. With respect to KO, our results coincide with Kerr (2008) (positive and significant effects) and, at least partially, also with Breschi et al. (2017).

4.4. Field and origin-country heterogeneity

Both inventor migration and the use of citations to acknowledge ideas diffusion, are highly heterogeneous across technological fields (WIPR, 2013). We explore this issue by dividing our inventor migration and citation flows across 5 technological fields, using the standard aggregation of IPC codes (Schmoch, 2008). As shown in Table 6, none of the five sectors shows a significant positive migration-diffusion relationship for the KI equation. In the meantime, the positive effects on KO show up in all domains, but they are especially strong in electrical and mechanical engineering.

[Table 6 about here]

Next, inspired by differences in STEM migration effects on different measures of globalization across types of countries (Kugler and Rapoport, 2007; Miguelez, 2018), we explore heterogeneous effects in the migration-diffusion relationship across the "North-North" and "South-North" axes. To do so, we multiply our main explanatory variable by two dummies: "High-income", valued 1 if the sending country is classified as a high-income country by the World Bank (before 2010), 0 otherwise, and "Middle/Low-income", valued 1 if the sending country is classified as a middle- or low-income country by the World Bank, 0 otherwise. We re-run regressions – again with country-pair FE, which are presented in Table 7: only inventors coming from middle- and low-income countries significantly affect KI. Conversely, migrant inventors are critical for KO in both cases (inventors from high- or middle- and low-income countries) – though the coefficient is significantly larger in favor of middle/low income countries.[9] It seems then that migrant inventors are more important for the "South-North" corridors than for the "North-North" ones. Together with results on interactions (Table 5), we interpret this as evidence of greater transaction costs in "South-North" country-pairs, as compared to "North-North" ones, due to larger cultural and technological differences. Indeed, the average technological proximity among developed-developed pairs is significantly greater (0.54) than between developing-developed countries (0.19). All this makes the role of STEM migrants especially relevant in these contexts. Besides, as already pointed out in Breschi et al. (2017), it could be that for the "North-North" corridors most of the knowledge travels within multinationals' boundaries, jumping between their different facilities located in different places, with migrant inventors playing a more nuanced role in this case. We address this particular point in the next subsection.

[Table 7 about here]

4.5. Multinationals and STEM migration

Table 8 reproduces the main regressions, removing intra-company citations – identified using the OECD HAN database, July 2014. As can be seen, results remain the same: they are not significant for knowledge inflows

---

[9] Wald tests indicate that differences in coefficients are significant in both cases.

(column 1), and are positive and significant for knowledge outflows (column 2). Columns 3 and 4 break the sample down into countries of origin, and again the results found in Table 7 are reproduced.

[Table 8 about here]

Next, we also re-compute our dependent variable using intra-company citations only. This is done in Table 9, where country-pair FE regressions are shown in columns 1 and 2. Interestingly, the coefficient for the relationship with KI now increases and becomes significant, which supports the idea that migrant inventors do bring in knowledge flows, but only within the boundaries of their firms. This is confirmed even more strongly when the focal variable is split between high-income and middle/low-income sending countries. As can be seen, now not only is the coefficient for middle/low-income countries positive and significant, but that of the high-income economies is too, confirming the idea that knowledge flows carried by migrants pass, at least partially, through multinationals at the same time. Again, the importance of STEM migration for intra-firm knowledge flows accords with findings by Oettl and Agrawal (2008) and Breschi et al. (2017).

[Table 9 about here]

4.6. Direct and indirect effects

We also explore differences in the migration-diffusion relationship, depending on whether citations occur among members of a given migrant community and their home colleagues (direct effects) or they include natives and migrants from other origins (indirect effects). We expect the former to show stronger effects, but the latter to influence diffusion, too. As we do not have information on nationality for all the inventors listed in the citing and cited patents, the results should be treated with care, but are still informative. Thus, as shown in Table 10, direct effects (both for KI and for KO) show positive coefficients, as expected. They are particularly large for KI, in fact. Results emerge also for indirect effects, which remain positive and significant. In fact, the coefficient for KI turns out to be significant when citations are split between direct and indirect effects.

[Table 10 about here]

4.7. Robustness analysis

We present some robustness checks in this section. As discussed in section 2 of the present paper, the large majority of empirical evidence on the relationship between STEM migration and knowledge diffusion concerns mainly the US, as it is, by far, the largest receiving talent country – especially from China and India, as well as the leading technology nation from which international spillovers emanate. In order to assess whether the results encountered in this paper, and in a large part of the related literature, can be extended beyond the US, we reproduce some of our regressions without the US as destination country.

This is done in Table 11. As in previous regressions, STEM migration does not impact KI, on average. Interestingly, the coefficient on KO is considerably reduced too, becoming non-significant, due to the importance of the US as an attractor of foreign talent as well as a source of knowledge and technology to all other nations. Columns 3 and 4 differentiate across countries of origin (high- vs middle/low-income countries), and again find that what matters for non-US countries is STEM migration from developing countries, as it shows a positive and significant relationship with both KI and KO. Thus, when removing the US as receiving country, it seems that STEM migrants are important only when differences across countries are more acute ("South-North" axis).

[Table 11 about here]

Table 12 removes the BRICS countries from the analysis (Brazil, Russia, India, China and South Africa). Results do not change to a large extent with respect to the baseline. However, the effect of migrants from middle-income

countries on KI, formerly positive and significant, does not arise this time. This is as expected, as a large majority of migrants from middle-income economies originate in BRICS countries.

[Table 12 about here]

Next, in Table 13 we focus on intra-European flows only. Columns 1 and 2 confirm that the migration-diffusion relationship is positive and significant, again, for KO, but not for KI (even slightly negative). Given that barely any flows come from non-high income countries, we focus instead on inter- and intra-company citations in columns 3 to 6. Results are repeated for inter-company citations (migrant inventors matter for KO). However, migrant inventors do matter for KI when these flows occur within the boundaries of the firm.

[Table 13 about here]

In further robustness checks, Table 14 shows the main results using only X, I and Y citations. These are particularly relevant documents, that may question the novelty and/or inventive step of using the citing patents (Jaffe and de Rassenfosse, 2017). This type of citation is possibly more important for inventors themselves, and less for examiners and lawyers (Criscuolo and Verspagen, 2008), which makes them more suitable for proxying knowledge flows. Fortunately, as shown in Table 14, our results and conclusions hold. In fact, the majority of coefficients are larger when using only "relevant documents".

[Table 14 about here]

5. Conclusion

In this paper we have used the gravity model to show how STEM migration – as measured by the number of migrant inventors – affects international knowledge diffusion – as measured by patent citations. While this research question is not new, systematic, global empirical evidence (especially beyond the US) is still scarce. Using inventors as a proxy for STEM migrants and their declared nationality to infer their migratory background, we have provided new results on the relationship between STEM migration and knowledge brought into their host countries, as well as knowledge sent back to their homelands. Further, we have also explored certain conditions for which these relationships are not linear: (1) when transaction costs are more acute ("South-North" axis); (2) within the boundaries of multinationals, and (3) beyond the US.

All in all, our results suggest that migrant inventors living in a given country are important for knowledge flows, not only to their homelands but also to their host countries – though to a lesser extent. Contrary to what has been advocated in the migration literature on the detrimental effect of high-skilled migration from low-income countries and an alarming brain drain, we find that low/middle-income countries benefit technologically from their migrant inventors living in high-income economies. At the same time, high-skilled migrants from low/middle-income countries also bring in some knowledge to their high-income host countries. Finally, we also learn that STEM migration and multinationals' strategies interact with each other with respect to knowledge diffusion, as well as the importance of the US in driving our results (with a few, notable exceptions).

Our research intends to convey the message that, instead of focusing the debate on brain drain issues, the attention of home and host countries' policymakers should be more oriented towards finding strategies that will establish and strengthen connections between STEM migrants and their non-mover peers, both at home and abroad, through adequate knowledge networks.

A few limitations are worth discussing. Our approach to STEM migration, based on inventors from the PCT with known nationality, could be an underestimate, as it misses inventors with a migratory background that have become nationals of their host country. If the likelihood of gaining citizenship differs across country-pairs, this could bias our estimates – though we build our explanatory variable using 1-year windows, so as to account for the most recent migrants only. Unfortunately, it is difficult to assess the severity of this potential bias. Related to this, PCT applications comprise only around 15-20% of all inventions worldwide. They are, indeed, only a

subsample of all patents (and all inventors). However, the underlying inventions are likely to have a larger economic and technological value than national applications (van Zeebroeck and van Pottelsberghe de la Potterie, 2011). Finally, as our analysis remains at the aggregate, country level (despite efforts to disentangle heterogeneous effects across countries), some particularities may remain hidden. More detailed, case study approaches would be required to uncover specific singularities.

References

Agrawal, A., Kapur, D., McHale, J., Oettl, A., 2011. Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. Journal of Urban Economics 69, 43–55. https://doi.org/10.1016/j.jue.2010.06.003

Alesina, A., Harnoss, J., Rapoport, H., 2016. Birthplace diversity and economic prosperity. J Econ Growth 21, 101–138. https://doi.org/10.1007/s10887-016-9127-6

Anderson, J.E., 2011. The gravity model. Annu. Rev. Econ. 3, 133–160.

Arora, A., Belenzon, S., Lee, H., 2018. Reversed citations and the localization of knowledge spillovers. J Econ Geogr 18, 495–521. https://doi.org/10.1093/jeg/lby015

Artuç, E., Docquier, F., Özden, Ç., Parsons, C., 2015. A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations. World Development, Migration and Development 65, 6–26. https://doi.org/10.1016/j.worlddev.2014.04.004

Audretsch, D.B., Feldman, M.P., 1996. R&D Spillovers and the Geography of Innovation and Production. American Economic Review 86, 630–40.

Bahar, D., Rapoport, H., 2018. Migration, Knowledge Diffusion and the Comparative Advantage of Nations. The Economic Journal 128, F273–F305. https://doi.org/10.1111/ecoj.12450

Baruffaldi, S.H., Landoni, P., 2012. Return mobility and scientific productivity of researchers working abroad: The role of home country linkages. Research Policy 41, 1655–1665.

Beine, M., Bertoli, S., Fernández-Huertas Moraga, J., 2016. A Practitioners' Guide to Gravity Models of International Migration. World Econ 39, 496–512. https://doi.org/10.1111/twec.12265

Beine, M., Docquier, F., Rapoport, H., 2001. Brain drain and economic growth: theory and evidence. Journal of Development Economics 64, 275–289. https://doi.org/10.1016/S0304-3878(00)00133-4

Belfanti, C.M., 2006. Between mercantilism and market: privileges for invention in early modern Europe. Journal of Institutional Economics 2, 319–338. https://doi.org/10.1017/S1744137406000439

Bhagwati, J., Hamada, K., 1974. The brain drain, international integration of markets for professionals and unemployment : A theoretical analysis. Journal of Development Economics 1, 19–42.

Borjas, G.J., Doran, K.B., 2015. Cognitive Mobility: Labor Market Responses to Supply Shocks in the Space of Ideas. Journal of Labor Economics 33, S109–S145. https://doi.org/10.1086/676659

Borjas, G.J., Doran, K.B., 2012. The Collapse of the Soviet Union and the Productivity of American Mathematicians*. The Quarterly Journal of Economics 127, 1143–1203. https://doi.org/10.1093/qje/qjs015

Bosetti, V., Cattaneo, C., Verdolini, E., 2015. Migration of skilled workers and innovation: A European Perspective. Journal of International Economics 96, 311–322. https://doi.org/10.1016/j.jinteco.2015.04.002

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

Breschi, S., Lissoni, F., 2005. Knowledge Networks from Patent Data, in: Moed, H.F., Glänzel, W., Schmoch, U. (Eds.), Handbook of Quantitative Science and Technology Research. Springer Netherlands, pp. 613–643.

Breschi, S., Lissoni, F., Miguelez, E., 2018. Return migrants' self-selection: Evidence for Indian inventors" – paper presented at the NBER conference on "The Role of Immigrants and

Foreign Students in Science, Innovation, and Entrepreneurship. National Bureau of Economic Research, Cambridge MA, April 27.

Breschi, S., Lissoni, F., Miguelez, E., 2017. Foreign-origin inventors in the USA: testing for diaspora and brain gain effects. J Econ Geogr 17, 1009–1038. https://doi.org/10.1093/jeg/lbw044

Burger, M., van Oort, F., Linders, G.-J., 2009. On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. Spatial Economic Analysis 4, 167–190. https://doi.org/10.1080/17421770902834327

Caligiuri, P., Bonache, J., 2016. Evolving and enduring challenges in global mobility. Journal of World Business 51, 127–141.

Cappelli, R., Montobbio, F., 2016. European integration and knowledge flows across European regions. Regional Studies 50, 709–727.

Chander, A., 2001. Diaspora Bonds'(2001). New York University Law Review 76, 1005–1043.

Chellaraj, G., Maskus, K.E., Mattoo, A., 2008. The Contribution of International Graduate Students to US Innovation. Review of International Economics 16, 444–462. https://doi.org/10.1111/j.1467-9396.2007.00714.x

Choudhury, P., 2016. Return migration and geography of innovation in MNEs: a natural experiment of knowledge production by local workers reporting to return migrants. J Econ Geogr 16, 585–610. https://doi.org/10.1093/jeg/lbv025

Choudhury, P., Kim, D.Y., 2019. The ethnic migrant inventor effect: Codification and recombination of knowledge across borders. Strategic Management Journal 40, 203–229. https://doi.org/10.1002/smj.2977

Cipolla, C.M., 1972. The diffusion of innovations in early modern Europe. Comparative Studies in Society and History 14, 46–52.

Clemens, M.A., 2014. Does development reduce migration. International Handbook on migration and Economic development 152–185.

Coe, D.T., Helpman, E., 1995. International R&D spillovers. European Economic Review 39, 859–887. https://doi.org/10.1016/0014-2921(94)00100-E

Coffano, M., Tarasconi, G., 2014. CRIOS - Patstat Database: Sources, Contents and Access Rules (SSRN Scholarly Paper No. ID 2404344). Social Science Research Network, Rochester, NY.

Criscuolo, P., Verspagen, B., 2008. Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. Research Policy, Special Section Knowledge Dynamics out of Balance: Knowledge Biased, Skewed and Unmatched 37, 1892–1908. https://doi.org/10.1016/j.respol.2008.07.011

Duguet, E., MacGarvie, M., 2005. How well do patent citations measure flows of technology? Evidence from French innovation surveys. Economics of Innovation and New Technology 14, 375–393. https://doi.org/10.1080/1043859042000307347

Eaton, J., Kortum, S., 1999. International technology diffusion: Theory and measurement. International Economic Review 40, 537–570.

Fassio, C., Montobbio, F., Venturini, A., 2019. Skilled migration and innovation in European industries. Research Policy 48, 706–718. https://doi.org/10.1016/j.respol.2018.11.002

Ferrucci, E., Lissoni, F., 2019. Foreign inventors in Europe and the United States: Diversity and Patent Quality. Research Policy.

Ganguli, I., 2015. Immigration and Ideas: What Did Russian Scientists "Bring" to the United States? Journal of Labor Economics 33, S257–S288. https://doi.org/10.1086/679741

Hedlund, G., 1986. The hypermodern MNC—a heterarchy? Human resource management 25, 9–35.

Hernandez, E., 2014. Finding a Home away from Home Effects of Immigrants on Firms' Foreign Location Choice and Performance. Administrative Science Quarterly 59, 73–108. https://doi.org/10.1177/0001839214523428

Hornung, E., 2014. Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia. The American Economic Review 104, 84–122. https://doi.org/10.1257/aer.104.1.84

Hunt, J., Gauthier-Loiselle, M., 2010. How Much Does Immigration Boost Innovation? American Economic Journal: Macroeconomics 2, 31–56. https://doi.org/10.1257/mac.2.2.31

Hymer, S.H., 1976. The international operations of national firms: A study of foreign direct investment.

Jaffe, A.B., 1986. Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value. The American Economic Review 76, 984–1001. https://doi.org/10.2307/1816464

Jaffe, A.B., de Rassenfosse, G., 2017. Patent citation data in social science research: Overview and best practices. Journal of the Association for Information Science and Technology 68, 1360–1374. https://doi.org/10.1002/asi.23731

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. The Quarterly Journal of Economics 108, 577–598. https://doi.org/10.2307/2118401

Kapur, D., McHale, J., 2005. Give us your best and brightest: The global hunt for talent and its impact on the developing world. Brookings Inst Press.

Keller, W., 2004. International technology diffusion. Journal of economic literature 42, 752–782.

Kemeny, T., Cooke, A., 2018. Spillovers from immigrant diversity in cities. J Econ Geogr 18, 213–245. https://doi.org/10.1093/jeg/lbx012

Kerr, S.P., Kerr, W., Özden, Ç., Parsons, C., 2016. Global Talent Flows. Journal of Economic Perspectives 30, 83–106. https://doi.org/10.1257/jep.30.4.83

Kerr, W.R., 2013. U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence (Working Paper No. 19377). National Bureau of Economic Research.

Kerr, W.R., 2008. Ethnic Scientific Communities and International Technology Diffusion. Review of Economics and Statistics 90, 518–537. https://doi.org/10.1162/rest.90.3.518

Kerr, W.R., Lincoln, W.F., 2010. The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention. Journal of Labor Economics 28, 473–508. https://doi.org/10.1086/651934

Kogut, B., Zander, U., 1993. Knowledge of the firm and the evolutionary theory of the multinational corporation. Journal of international business studies 24, 625–645.

Kugler, M., Levintal, O., Rapoport, H., 2018. Migration and Cross-Border Financial Flows. World Bank Econ Rev 32, 148–162. https://doi.org/10.1093/wber/lhx007

Kugler, M., Rapoport, H., 2007. International labor and capital flows: Complements or substitutes? Economics Letters 94, 155–162. https://doi.org/10.1016/j.econlet.2006.06.023

Kuznetsov, Y., 2006. Diaspora Networks and the International Migration of Skills : How Countries Can Draw on their Talent Abroad (World Bank Publications). The World Bank.

Lissoni, F., 2018. International migration and innovation diffusion: an eclectic survey. Regional Studies 52, 702–714.

Luu, L., 2005. Immigrants and the Industries of London, 1500-1700. Ashgate Aldershot.

MacGarvie, M., 2005. The determinants of international knowledge diffusion as measured by patent citations. Economics Letters 87, 121–126. https://doi.org/10.1016/j.econlet.2004.09.011

MacKinnon, J.G., Magee, L., 1990. Transforming the Dependent Variable in Regression Models. International Economic Review 31, 315–339. https://doi.org/10.2307/2526842

Mayer, T., Zignago, S., 2011. Notes on CEPII's Distances Measures: The GeoDist Database (SSRN Scholarly Paper No. ID 1994531). Social Science Research Network, Rochester, NY.

Meyer, J.-B., 2001. Network Approach versus Brain Drain: Lessons from the Diaspora. International Migration 39, 91–110. https://doi.org/10.1111/1468-2435.00173

Meyer, J.-B., Brown, M., 1999. Scientific diasporas: a new approach to the brain drain. Discussion paper prepared for the World Conference on Science, UNESCO-ICSU, Budapest, 26 June-1 July 1999 2002.

Miguelez, E., 2018. Inventor Diasporas and the Internationalization of Technology. World Bank Economic Review 32, 41–63. https://doi.org/10.1093/wber/lhw013

Miguelez, E., Fink, C., 2013. Measuring the International Mobility of Inventors: A New Database. World Intellectual Property Organization - Economics and Statistics Division.

Minbaeva, D.B., Michailova, S., 2004. Knowledge transfer and expatriation in multinational corporations: The role of disseminative capacity. Employee relations 26, 663–679.

Moretti, E., 2004. Human capital externalities in cities, in: Handbook of Regional and Urban Economics. Elsevier, pp. 2243–2291.

Moser, P., Voena, A., Waldinger, F., 2014. German Jewish Émigrés and US Invention. American Economic Review 104, 3222–3255. https://doi.org/10.1257/aer.104.10.3222

Nanda, R., Khanna, T., 2010. Diasporas and domestic entrepreneurs: Evidence from the Indian software industry. Journal of Economics & Management Strategy 19, 991–1012.

Nathan, M., 2015. Same difference? Minority ethnic inventors, diversity and innovation in the UK. J Econ Geogr 15, 129–168. https://doi.org/10.1093/jeg/lbu006

Nelson, R.R., Phelps, E.S., 1966. Investment in humans, technological diffusion, and economic growth. The American economic review 56, 69–75.

Oettl, A., Agrawal, A., 2008. International labor mobility and knowledge flow externalities. Journal of International Business Studies 39, 1242–1260. https://doi.org/10.1057/palgrave.jibs.8400358

Peri, G., 2005. Determinants of Knowledge Flows and Their Effect on Innovation. Review of Economics and Statistics 87, 308–322. https://doi.org/10.1162/0034653053970258

Polanyi, M., 1958. Tacit knowledge: Toward a post-critical philosophy. Chicago: University of.

Rayp, G., Ruyssen, I., Standaert, S., 2017. Measuring and Explaining Cross-Country Immigration Policies. World Development 95, 141–163.

Santos Silva, J.M.C., Tenreyro, S., 2006. The Log of Gravity. Review of Economics and Statistics 88, 641–658. https://doi.org/10.1162/rest.88.4.641

Saxenian, A., 2006. The New Argonauts: Regional Advantage in a Global Economy. Harvard University Press.

Saxenian, A., Motoyama, Y., Quan, X., 2002. Local and global networks of immigrant professionals in Silicon Valley. Public Policy Institute of California, San Francisco, CA.

Schmoch, U., 2008. Concept of a technology classification for country comparisons. Final report to the World Intellectual Property Organization (WIPO), Fraunhofer Institute for Systems and Innovation Research, Karlsruhe.

Singh, J., 2008. Distributed R&D, cross-regional knowledge integration and quality of innovative output. Research Policy 37, 77–96. https://doi.org/10.1016/j.respol.2007.09.004

Sorenson, O., Rivkin, J.W., Fleming, L., 2006. Complexity, networks and knowledge flow. Research policy 35, 994–1017.

Stephan, P.E., Levin, S.G., 2001. Exceptional contributions to US science by the foreign-born and foreign-educated. Population Research and Policy Review 20, 59–79.

Storper, M., Venables, A.J., 2004. Buzz: face-to-face contact and the urban economy. J Econ Geogr 4, 351–370. https://doi.org/10.1093/jnlecg/lbh027

Teece, D.J., 1977. Technology transfer by multinational firms: The resource cost of transferring technological know-how. The economic journal 87, 242–261.

Thompson, P., 2006. Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations. Review of Economics and Statistics 88, 383–388. https://doi.org/10.1162/rest.88.2.383

Thompson, P., Fox-Kean, M., 2005. Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. The American Economic Review 95, 450–460.

Trippl, M., 2013. Scientific Mobility and Knowledge Transfer at the Interregional and Intraregional Level. Regional Studies 47, 1653–1667. https://doi.org/10.1080/00343404.2010.549119

UN-DESA and OECD, 2013. World Migration in Figures.

Useche, D., Miguelez, E., Lissoni, F., 2019. Highly skilled and well connected: Migrant inventors in cross-border M&As. J Int Bus Stud. https://doi.org/10.1057/s41267-018-0203-3

van Zeebroeck, N., van Pottelsberghe de la Potterie, B., 2011. Filing strategies and patent value. Economics of Innovation and New Technology 20, 539–561. https://doi.org/10.1080/10438591003668646

Vandenbussche, J., Aghion, P., Meghir, C., 2006. Growth, distance to frontier and composition of human capital. Journal of economic growth 11, 97–127.

Winters, J.V., 2014. STEM graduates, human capital externalities, and wages in the US. Regional Science and Urban Economics 48, 190–198.

Table 1. Citations and migration corridors, 2006-2010

Citing country	Cited country	# of citations	Citing country	Cited country	# of citations
Top-10 corridors			**Top 10-20 corridors**		
Germany	US	90925.26	China	US	21979.27
US	Japan	78024.24	Israel	US	19767.45
UK	US	47207.88	R. Korea	US	19754.45
US	Germany	44903.69	US	R. Korea	19662.8
France	US	40987.69	US	Canada	19258.18
Germany	Japan	39445.01	Australia	US	18999.21
US	UK	28822.09	Italy	US	18831.14
Canada	US	28651.24	Sweden	US	18810.68
Japan	US	25303.78	US	France	15926.44
Netherlands	US	23852.72	France	Germany	14966.33

Citing country	Cited country	# of citations	Citing country	Cited country	# of citations
Top-10 inter-company corridors			**Top-10 intra-company corridors**		
Germany	US	87853.57	Germany	US	2949.833
US	Japan	75664.39	US	Japan	2272.237
UK	US	45119.2	US	Germany	2130.503
US	Germany	42722.13	UK	US	2033.256
France	US	39759.93	US	UK	1675.753
Germany	Japan	38925.94	Japan	US	1426.976
Canada	US	27419.32	Canada	US	1218.286
US	UK	27106.1	France	US	1160.141
Japan	US	23855.05	Netherlands	US	992.3494
Netherlands	US	22828.07	China	US	984.137

Origin country	Dest. country	# inventors	Origin country	Dest. country	# inventors
Top-10 corridors			**Top 10-20 corridors**		
China	US	27,696	France	Switzerland	1,880
India	US	21,712	Israel	US	1,878
Canada	US	11,364	Australia	US	1,783
UK	US	8,313	Netherlands	US	1,670
Germany	US	5,895	France	Germany	1,492
Germany	Switzerland	4,952	Italy	US	1,492
R. Korea	US	4,877	China	Japan	1,463
France	US	3,898	Germany	Netherlands	1,335
Japan	US	2,844	Austria	Germany	1,308
Russia	US	2,309	Turkey	US	1,233

Source: OECD Citations database from July 2014 and Miguelez and Fink (2013).

Table 2. Baseline results: Does inventor migration affect knowledge flows?

	(1) KI	(2) KO	(3) KI	(4) KO
Migrant inventors	0.0844***	0.0867***	0.0503***	0.0537***
	(0.0126)	(0.0159)	(0.0109)	(0.0148)
Contiguity	0.0268	0.0259	-0.0117	0.0111
	(0.0542)	(0.0571)	(0.0330)	(0.0362)
Colony	0.0213	0.0134	0.0210	0.00887
	(0.0350)	(0.0328)	(0.0226)	(0.0287)
Common official language	0.243***	0.239***	0.114***	0.121***
	(0.0430)	(0.0479)	(0.0309)	(0.0324)
ln(Distance b/ capitals)	-0.0949***	-0.0924***	-0.0407***	-0.0542***
	(0.0211)	(0.0243)	(0.0136)	(0.0172)
# inventors in i	0.595***	0.546***	0.594***	0.583***
	(0.0708)	(0.102)	(0.0714)	(0.103)
# inventors in j	0.560***	0.657***	0.589***	0.687***
	(0.0808)	(0.0398)	(0.0820)	(0.0419)
Average citations in i	-0.378***	0.208***	-0.362***	0.258***
	(0.123)	(0.0702)	(0.108)	(0.0741)
Average citations in j	0.430***	-0.331**	0.467***	-0.369***
	(0.140)	(0.146)	(0.134)	(0.132)
Exports			0.375**	-0.227
			(0.148)	(0.141)
Imports			0.00740	0.323**
			(0.190)	(0.136)
Stocks migrants 2000			0.0478**	0.0740***
			(0.0220)	(0.0198)
Tecnological similarity			1.651***	1.717***
			(0.126)	(0.144)
Constant	-2.860**	-3.310**	-5.786***	-6.152***
	(1.217)	(1.312)	(1.343)	(1.405)
Observations	88,635	88,635	88,635	88,635
Origin country FE	Yes	Yes	Yes	Yes
Destination country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The reason why observations are less than the implied 33 receiving countries * (133-1) sending * 21 years is that inventor migration data are not existent for some country-year pairs (e.g., ex-USSR republics before 1993). All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for dummies, the index of technological similarity and distance between capitals, which is log transformed.

Table 3. Instrumental variables regression. Instrument: Migration policy index

	(1) Migrant inventors	(2) Equivalent baseline KI	(3) KO	(4) IV KI	(5) KO
Migrant inventors		0.0683***	0.0426**	0.0583	0.322***
		(0.0137)	(0.0179)	(0.0433)	(0.0600)
Mig. Policy Index	0.104***				
	(0.0251)				
Controls	Yes	Yes	Yes	Yes	Yes
1st stage F-stat	22.17				
p-value	0.000				
Observations	39,600	39,600	39,600	39,600	39,600
Origin & Dest. FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Controls in all regressions include Contiguity, Colony, Common official language, ln(Distance b/ capitals), # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, Stocks migrants 2000, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for dummies, the index of technological similarity, the migration policy index, and distance between capitals, which is log transformed.

Table 4. Baseline with country-pair FE

	(1) KI	(2) KO
Migrant inventors	-0.00895	0.0543***
	(0.0145)	(0.0198)
Controls	Yes	Yes
Observations	50,991	51,926
Origin & dest. FE	No	No
Year FE	Yes	Yes
Country-pair FE	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 5. Interactions

	Migrant inventors* Tech.similarity	Migrant inventors* Contiguity	Migrant inventors* Colony	Migrant inventors* Language	Migrant inventors* Geo Proximity
KI	0.0442	0.0701*	-0.0154	-0.0181	-0.0409***
	(0.0738)	(0.0379)	(0.0493)	(0.0348)	(0.0144)
KO	-0.120**	0.0707*	-0.00723	-0.0207	-0.0378**
	(0.0488)	(0.0378)	(0.0358)	(0.0300)	(0.0171)

Notes: Country-pair level clustered standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 6. Field heterogeneity

	(1) Electrical engineering	(2) Instruments	(3) Chemistry	(4) Mechanical engineering	(5) Other fields
KI	0.0113	-0.0351***	-0.0195	-0.0149	-0.0149
	(0.0137)	(0.0129)	(0.0130)	(0.0116)	(0.0116)
KO	0.0420*	0.0299**	0.0266**	0.0413***	0.0244***
	(0.0251)	(0.0130)	(0.0105)	(0.0111)	(0.00886)

Notes: Country-pair level clustered standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 7. Split inventor migration between high and medium/low-income countries

	(1) KI	(2) KO
Migrant inventors *High-income	-0.0128	0.0470**
	(0.0153)	(0.0199)
Migrant inventors *Middle/Low-income	0.135***	0.149***
	(0.0468)	(0.0301)
Wald test	7.67	12.89
Prob.	0.0056	0.0003
Observations	50,991	51,926
Origin & destination country FE	No	No
Year FE	Yes	Yes
Country-pair FE	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 8. Only inter-company citations

	(1) KI	(2) KO	(3) KI	(4) KO
Migrant inventors	-0.0117	0.0544***		
	(0.0149)	(0.0202)		
Migrant inventors *High-income			-0.0154	0.0471**
			(0.0157)	(0.0203)
Migrant inventors *Middle/Low-income			0.130***	0.149***
			(0.0464)	(0.0304)
Controls	Yes	Yes	Yes	Yes
Wald test			7.46	12.66
Prob.			0.0063	0.0004
Observations	50,478	51,577	50,478	51,577
Origin & destination country FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 9. Only intra-company citations

	(1) KI	(2) KO	(3) KI	(4) KO
Migrant inventors	0.0900***	0.0549***		
	(0.0258)	(0.0199)		
Migrant inventors High-income			0.0833***	0.0494**
			(0.0262)	(0.0209)
Migrant inventors *Middle/Low-income			0.257**	0.146***
			(0.101)	(0.0547)
Controls	Yes	Yes	Yes	Yes
Wald test			2.95	2.86
Prob.			0.0859	0.0905
Observations	22,660	23,520	22,660	23,520
Origin country FE	No	No	No	No
Destination country FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 10. Direct and indirect effects of diaspora networks on knowledge flows

	(1) KI Direct	(2) KI Indirect	(3) KO Direct	(4) KO Indirect
Migrant inventors	0.354***	0.0648**	0.0903***	0.0565***
	(0.0496)	(0.0269)	(0.0271)	(0.0190)
Observations	27,026	49,353	24,644	51,079
Controls	Yes	Yes	Yes	Yes
Origin & dest. country FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 11. Excluding the US

	(1) KI	(2) KO	(3) KI	(4) KO
Migrant inventors	0.00191	0.0250		
	(0.0116)	(0.0187)		
Migrant inventors *High-income			-0.00264	0.0201
			(0.0119)	(0.0191)
Migrant inventors*Middle/Low-income			0.218***	0.100***
			(0.0322)	(0.0259)
Controls	Yes	Yes	Yes	Yes
Wald test			38.21	8.48
Prob.			0.0000	0.0036
Observations	48,367	49,260	48,367	49,260
Origin & dest. country FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 12. Removes BRICS

	(1) KI	(2) KO	(3) KI	(4) KO
Migrant inventors	-0.0127	0.0628***		
	(0.0173)	(0.0203)		
Migrant inventors*High-income			-0.0131	0.0613***
			(0.0175)	(0.0204)
Migrant inventors*Middle/Low-income			0.0311	0.142***
			(0.0399)	(0.0334)
Controls	Yes	Yes	Yes	Yes
Wald test			0.92	6.99
Prob.			0.3370	0.0082
Observations	47,705	48,598	47,705	48,598
Origin & dest. country FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 13. Keeps only European countries as origin and destination

	(1)	(2)	(3)	(4)	(5)	(6)
	KI	KO	KI	KO	KI	KO
	All flows		Inter-company flows		Intra-company flows	
Migrant inventors	-0.0231*	0.0684***	-0.0268**	0.0681***	0.0895**	0.0737**
	(0.0124)	(0.0138)	(0.0123)	(0.0140)	(0.0383)	(0.0344)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,623	14,312	13,543	14,192	7,803	7,945
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Table 14. Only X, I, Y citations used

	(1)	(2)	(3)	(4)
	KI	KO	KI	KO
Migrant inventors	0.00403	0.0588***		
	(0.0152)	(0.0193)		
Migrant inventors *High-income			-0.000183	0.0505***
			(0.0159)	(0.0195)
Migrant inventors *Middle/Low-income			0.156***	0.184***
			(0.0442)	(0.0231)
Controls	Yes	Yes	Yes	Yes
Wald test			10.03	32.28
Prob.			0.0015	0.0000
Observations	44,029	46,091	44,029	46,091
Origin & destination country FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Country-pair FE	Yes	Yes	Yes	Yes

Notes: Country-pair level clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Observations decrease with respect to table 2, due to the fact that non-linear models (e.g., Poisson) removes country-pair observations in the absence of time variation (all zero outcomes) if country-pair FE are included. Controls in all regressions include # inventors in i, # inventors in j, Average citations in i, Average citations in j, Exports, Imports, and Technological similarity. All explanatory variables are transformed using the inverse hyperbolic sine transformation (MacKinnon and Magee, 1990), except for the index of technological similarity.

Inventor Migration and Knowledge Flows: A Two-Way Communication Channel?

APPENDIXES
(for on-line publication only)

Content

Appendix A1: Descriptive statistics

Table A1.1. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
KI	88,635	30.43	359.22	0	20740.89
KO	88,635	31.93	361.79	0	20740.89
KI inter-firm	88,635	29.38	347.57	0	19889.66
KO inter-firm	88,635	30.84	350.02	0	19889.66
KI intra-firm	88,635	1.05	12.62	0	851.24
KO intra-firm	88,635	1.09	12.74	0	851.24
KI ethnic	88,635	1.41	20.15	0	1249.20
KO ethnic	88,635	0.70	10.03	0	743.51
KI not ethnic	88,635	24.39	295.56	0	19194.96
KO not ethnic	88,635	29.17	327.01	0	17502.39
Migrant inventors	88,635	0.36	0.99	0	9.32
Exports	88,635	13.11	8.43	0	27.28
Imports	88,635	14.12	8.22	0	27.23
Stock migrants 2000	88,635	4.64	3.89	0	15.49
Technological similarity	88,635	0.31	0.29	0	1
Contiguity	88,635	0.02	0.15	0	1
Colony	88,635	0.03	0.17	0	1
Common official language	88,635	0.10	0.29	0	1
ln(distance)	88,635	8.51	0.90	4.09	9.88
# of inventors in i	88,635	7.54	2.39	0	12.33
# of inventors in j	88,635	3.07	3.28	0	12.33
Average citations in i	88,635	0.88	0.67	0	4.61
Average citations in j	88,635	1.33	0.29	0	2.64

Notes: All explanatory variables are transformed using the inverse hyperbolic sine transformation, except for dummies, the index of technological similarity, the migration policy index, and distance between capitals, which is log transformed.

Table A1.2. Correlation matrix

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	KI	1														
2	KO	0.7	1													
3	Migrant inventors	0.3	0.4	1												
4	Exports	0.1	0.1	0.4	1											
5	Imports	0.1	0.1	0.3	0.9	1										
6	Stock migrants 2000	0.1	0.1	0.5	0.4	0.4	1									
7	Tech. similarity	0.2	0.2	0.4	0.6	0.5	0.4	1								
8	Contiguity	0.1	0.1	0.2	0.1	0.1	0.2	0.2	1							
9	Colony	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	1						
10	Common language	0.0	0.0	0.1	0.0	0.0	0.2	0.0	0.1	0.3	1					
11	ln(distance)	0.0	0.0	-0.2	-0.2	-0.2	-0.2	-0.2	-0.4	-0.1	0.0	1				
12	# of inventors in i	0.1	0.1	0.4	0.3	0.3	0.4	0.1	0.0	0.1	0.1	0.0	1			
13	# of inventors in j	0.2	0.2	0.4	0.6	0.5	0.4	0.8	0.2	0.0	0.0	-0.3	0.0	1		
14	Average citations in i	0.0	0.0	0.2	0.4	0.3	0.2	0.5	0.1	0.0	0.0	-0.1	0.0	0.5	1	
15	Average citations in j	0.0	0.0	0.0	-0.1	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	-0.1	0.2	1

Notes: All explanatory variables are transformed using the inverse hyperbolic sine transformation, except for dummies, the index of technological similarity, the migration policy index, and distance between capitals, which is log transformed.

Appendix A2: Citation and migration corridors originating from developing countries

Table A.2.1. Citations corridors – without high-income as citing countries (total flow of citations for the top-20 country-pairs for the period 2006-2010)

Citing country	Cited country	Total share of citations	Citation share Cum. (%)
China	US	22,055.15	20.07
China	Japan	8,533.97	27.83
India	US	7,673.99	34.81
China	South Korea	5,111.61	39.46
China	Germany	4,143.47	43.23
Russian Federation	US	3,281.55	46.22
South Africa	US	2,885.64	48.85
Brazil	US	2,874.61	51.47
China	UK	2,078.99	53.36
India	Japan	1,747.44	54.95
Mexico	US	1,737.97	56.53
China	France	1,654.54	58.04
China	Finland	1,594.19	59.49
Turkey	US	1,451.02	60.81
India	Germany	1,392.88	62.08
China	Canada	1,384.37	63.34
Malaysia	US	1,193.37	64,43
China	Sweden	1,162.83	65.49
Turkey	Germany	1,019.76	66.42
India	UK	965.89	67.3

Source: OECD citations database, July 2014

Table A.2.2 Migration corridors – without high-income origin countries (total of inventor immigrants for the top-20 country pairs for the period 2006-2010)

Origin country	Destination country	Total migration	Cum. Migration share (%)
China	US	27,696	31.79
India	US	21,712	56.71
Russian Federation	US	2,309	59.36
China	Japan	1,463	61.04
Turkey	US	1,233	62.46
China	Singapore	1,149	63.78
Iran	US	960	64.88
Brazil	US	763	65.75
Mexico	US	722	66.58
Romania	US	710	67.39
Russian Federation	Germany	702	68.2
India	Singapore	610	68.9
Malaysia	Singapore	607	69.6
Ukraine	US	601	70.29
China	Germany	555	70.92
China	UK	545	71.55
Malaysia	US	484	72.11
Argentina	US	478	72.65
South Africa	US	414	73.13
India	UK	393	73.58

Source: Miguelez and Fink (2013).

Appendix A3: List of countries included in the analysis

Table A.3.1. Countries included in the analysis

Receiving countries	Sending countries			
Australia	Albania	El Salvador	Lebanon	Senegal
Austria	Algeria	Equatorial Guinea	Liberia	Sierra Leone
Belgium	Argentina	Estonia	Libya	Singapore
Canada	Armenia	Ethiopia	Lithuania	Slovakia
Chile	Australia	Finland	Luxembourg	Slovenia
Czech Republic	Austria	France	Madagascar	South Africa
Denmark	Azerbaijan	Gabon	Malawi	Spain
Estonia	Bahamas	Gambia	Malaysia	Sri Lanka
Finland	Bangladesh	Georgia	Mali	Sudan
France	Belarus	Germany	Malta	Suriname
Germany	Belgium	Ghana	Mauritania	Swaziland
Greece	Bolivia	Greece	Mexico	Sweden
Hungary	Bosnia and Herzegovina	Guatemala	Mongolia	Switzerland
Iceland	Botswana	Guinea	Morocco	Syria
Ireland	Brazil	Haiti	Namibia	Macedonia
Israel	Bulgaria	Honduras	Netherlands	Tajikistan
Italy	Burkina Faso	Hungary	New Zealand	Thailand
Japan	Cameroon	Iceland	Nicaragua	Trinidad and Tobago
Luxembourg	Canada	India	Niger	Tunisia
Mexico	Chad	Indonesia	Nigeria	Turkey
Netherlands	Chile	Iran	Norway	Uganda
New Zealand	China	Iraq	Oman	Ukraine
Norway	Colombia	Ireland	Pakistan	UAE
Poland	Congo	Israel	Panama	UK
Portugal	Costa Rica	Italy	Paraguay	Tanzania
R. of Korea	Croatia	Jamaica	Peru	US
Slovakia	Cuba	Japan	Philippines	Uruguay
Slovenia	Cyprus	Jordan	Poland	Uzbekistan
Spain	Czech Republic	Kazakhstan	Portugal	Venezuela
Sweden	Côte d'Ivoire	Kenya	Qatar	Viet Nam
Switzerland	Denmark	Kuwait	R. of Korea	Zambia
UK	Dominican Rep.	Kyrgyzstan	R. of Moldova	Zimbabwe
US	Ecuador	Latvia	Russian Fed.	
	Egypt		Saudi Arabia	