Abstract
Multinational enterprises (MNEs) geographically distribute their research and development (R&D) activities to source foreign knowledge that is critical to feed the innovation process with valuable inputs developed in different locations, as well as to adapt their technology to new markets. However, literature suggests that, in spite of the establishment of foreign R&D subsidiaries in selected host-countries, MNEs face serious challenges in accomplishing these tasks. Building on international business and migration studies, we argue that ethnic inventors play a role in alleviating these challenges, by acting as a bridge between their countries of origin (CoO) and the MNE headquarters. Using USPTO data on global collaborative patents assigned to US-based MNEs operating in knowledge-intensive industries over the period 1975-2009, our empirical analysis shows that ethnic inventors from a given CoO foster the integration of knowledge emanating from that CoO in the intra-corporate innovation process. Further analyses demonstrate that this effect is bounded to knowledge developed within the MNE, confirming the superior role of firms compared to external market mechanisms in the transfer of knowledge across borders. Finally, contrary to our theoretical predictions suggesting that ethnic inventors also play a role in facilitating the adaptation and exploitation of MNE’s technology in their CoO market, we find no significant evidence of such an effect.
Ties that bind: the role of ethnic inventors in multinational enterprises’ knowledge creation

Abstract

Multinational enterprises (MNEs) geographically distribute their research and development (R&D) activities to source foreign knowledge that is critical to feed the innovation process with valuable inputs developed in different locations, as well as to adapt their technology to new markets. However, literature suggests that, in spite of the establishment of foreign R&D subsidiaries in selected host-countries, MNEs face serious challenges in accomplishing these tasks. Building on international business and migration studies, we argue that ethnic inventors play a role in alleviating these challenges, by acting as a bridge between their countries of origin (CoO) and the MNE headquarters. Using USPTO data on global collaborative patents assigned to US-based MNEs operating in knowledge-intensive industries over the period 1975-2009, our empirical analysis shows that ethnic inventors from a given CoO foster the integration of knowledge emanating from that CoO in the intra-corporate innovation process. Further analyses demonstrate that this effect is bounded to knowledge developed within the MNE, confirming the superior role of firms compared to external market mechanisms in the transfer of knowledge across borders. Finally, contrary to our theoretical predictions suggesting that ethnic inventors also play a role in facilitating the adaptation and exploitation of MNE’s technology in their CoO market, we find no significant evidence of such an effect.

Keywords: migration, inventors, multinational enterprises, knowledge integration, knowledge exploitation, ethnic ties

JEL codes: F22, F23, O15, O32
Introduction

The global mass migration of skilled workers represents a huge trend in today’s economy (OECD, 2017). The overall number of high-skilled migrants to OECD countries increased from 12 million in 1990 to 28 million in 2010 (Kerr et al., 2016), and comprises an increasing percentage of total migrant flows in all these countries except for the US (Kerr et al., 2017). Data releases by the US Citizenship and Immigration Services (USCIS) department highlight that, in the last years, applications for H-1B work visas in the United States (US) have systematically exceeded the program’s available approvals. These high skill migrants are also arriving from a wider range of origin countries, a condition that seems to yield development benefits (Alesina, et al., 2016).

This phenomenon has generated debate about the policies that both origin and destination countries should adopt to regulate this flow of valuable human capital. The availability of migrant skilled workers coupled with the shortage of domestic human capital in many advanced economies (Boeri et al., 2012) has also changed managerial practices related to the global sourcing of talent. Multinational enterprises (MNEs) are a major factor driving high skilled migration as they increasingly redeploy talent hired in one country across their global network of subsidiaries (Lewin et al., 2009).

The outcome of this trend is evident in patent data, which show that a substantial share of patents assigned to US companies have been developed by teams involving ethnic inventors, i.e. inventors based in the MNE home-country (in this case, the US) and characterized by a foreign ethnicity (Kerr and Kerr, 2018). Recent literature has also documented that such inventors are disproportionately involved in particular types of global collaborative patents1, referred to as own-ethnicity collaborative patents, which “exhibit a specific match between the ethnicity of the US-

---

1 Following Kerr and Kerr (2018), global collaborative patents are defined as patents involving at least one inventor within the US and at least one inventor outside of the US.
based inventor and the foreign region in which the other members of the inventor team are located” (Kerr and Kerr, 2018: 2). Further, when these skilled workers are deployed in the country of their ethnicity, “local employees with returnee managers file disproportionately more US patents” (Choudhury, 2016: 585). Based on this evidence, it seems relevant to explore the following research question: How do own-ethnic inventors\(^2\) contribute to intra-MNE knowledge creation? International Business (IB) literature suggests that MNEs geographically distribute their R&D activities both to source foreign knowledge that is critical to feed the innovation process with valuable inputs from different locations (knowledge integration), and to exploit their technological assets in new geographical markets (knowledge exploitation) (Kuemmerle, 1999; Cantwell and Mudambi, 2005). However, the accomplishment of these tasks is neither immediate nor effortless. For instance, it has been demonstrated that MNEs face serious challenges in integrating foreign technology (Teece, 1981; Gupta and Govindarajan, 2000) because the transfer and assimilation of sticky and causally ambiguous knowledge across long distances is difficult even within the internal organizational network (Polanyi, 1966; Szulanski, 1996; Zander and Kogut, 1995). Similarly, the exploitation of MNE knowledge in a foreign market is based on adaptation processes that require knowledge of the foreign market, which is unlikely to be gained without a sufficient degree of embeddedness in the foreign country’s institutional context (Meyer et al., 2011).

Building on migration studies (Meyer and Brown, 1999; Meyer, 2001; Levitt and Jaworski, 2007; Docquier and Rapoport, 2012; Nowicka, 2014), we argue that own-ethnic inventors could serve a bridging function between their country of origin (CoO) and the MNE headquarters. There is some evidence that “return migrants act as a ‘bridge’ to transfer knowledge from the MNE headquarters to the local employees working for them” (Choudhury, 201: 585), but this evidence is drawn from

---

\(^2\) Following Kerr and Kerr (2018), own-ethnic inventors are defined as US-based inventors whose foreign-ethnicity matches the foreign region in which the other members of the inventor team are located. (Kerr and Kerr, 2018)
“a relatively small sample [...] in a single-firm setting” (Choudhury, 2016: 608) leading to potential biases and limited generalizability. We extend this finding by investigating this potential bridging effect in the opposite direction, i.e. from the own-ethnic inventor’s CoO to the MNE headquarters, using a broad sample of global collaborative patents. Furthermore, we explore this bridging function both in the context of knowledge integration and knowledge exploitation. Specifically, we posit that thanks to the unique ties that bind them to their CoO, own-ethnic inventors can facilitate knowledge integration from their CoO, thereby improving access to, and re-use of valuable knowledge that is deeply embedded in the CoO. In other words, they serve to ameliorate the well-known barriers to the transfer of sticky knowledge (e.g., von Hippel, 1994; Szulanski, 1996). Moreover, we investigate whether own-ethnic inventors can promote MNE knowledge exploitation in the CoO market. Previous literature has widely documented that MNEs devote significant R&D efforts to adapt particular products, processes, functions or procedures to foreign market conditions to exploit their technological assets across borders, either externally or internally (i.e., for instance by supplying new technology to their own production facilities abroad) (Dunning and Lundan, 2009). Due to own-ethnic inventors’ greater exposure to the CoO’s habits, industry norms, institutions and general business environment (D’Ambrosio et al., 2018), they are likely to enjoy a privileged position to provide insights into how to effectively apply and adapt the MNE technological knowledge to the CoO. Thus, we expect that new technology developed by global collaborative teams involving own-ethnic inventors should be easier to redeploy in the CoO market.

Using USPTO data on global collaborative patents assigned to US-based MNEs operating in knowledge-intensive industries over the period 1975-2009, our empirical analysis shows that own-ethnic inventors from a given CoO foster the integration of knowledge emanating from that CoO
in the intra-corporate context. Further analyses demonstrate that this effect is bounded to CoO knowledge that has been developed within the MNE and does not extend to knowledge developed in the CoO but outside the MNE’s organizational boundaries, confirming the superior role of firms compared to external market mechanisms in the transfer of knowledge across borders. Conversely, the empirical analysis does not support the knowledge exploitation role of own-ethnic inventors. This seems to suggest that own-ethnic inventors are effective carriers of knowledge, but not of business or market information (Cowan et al., 2000). This finding provides some preliminary insights into the recent debate on the content of diffusion flows activated by highly skilled migrants (Lissoni, 2018).

More generally, our study contributes to the nascent stream of literature on the contribution of highly skilled migrants to the knowledge creation processes of MNEs (Choudhury, 2016). In doing so, it answers to the recent call for greater consideration of the role of firms in studies on migration (Kerr et al., 2015), a role that future research should continue to explore given that firms, and particularly MNEs, are key actors in determining and directing highly skilled migration flows.

**Background and prior literature**

*High-skill migration*

Human migration dates back to the dawn of humanity, but modern large-scale migration mainly dates to the last two centuries. Nowadays, the destinations for virtually all high skill migrants³ are OECD countries (boxes 1 and 2 in Table 1). As the economies of most destination countries became more technology- and knowledge-intensive, their demands for labor changed accordingly. By the late twentieth century, virtually all OECD countries faced shortages of high skill labor and

³ We define high-skilled workers as “those having a university degree or extensive/equivalent experience in a given field” (Iredale, 2001: 8).
they put in place policies for attracting highly-skilled migrants (Doomernik, et al., 2009): employment-based, occupation-driven and human-capital oriented (Koslowski, 2014). The former policies allow migrants to remain in the destination country if they are able to secure and provide documentation of appropriate work status. Occupation-driven policies admit migrants who are qualified in occupations that the government decides are in short supply. Finally, human-capital oriented policies screen migrants based on their skills and qualifications, typically using these to award points. These policies affect the international distribution of high skill migrants across countries, thus favoring higher concentrations in specific occupations. For example, foreign-born individuals made up 27 percent of all physicians and surgeons and more than 35 percent of current medical residents in the United States in 2010 (Kerr et al., 2017).

The migration literature recognizes that the movement of individuals across national borders may be temporary or permanent. In some cases, like wars or famines, migrants may wish to return to their home countries when the crisis is past (box 3 in Table 1). However, this literature still views the contribution of highly skilled migrants (boxes 1 and 2 in Table 1) to their destination economy as the returns from the direct local application of their human capital (Sassen, 1990).

*MNEs, national systems of innovation, and highly skilled migrants*

Over the last several decades, MNEs have been increasingly dispersing their innovation activities around the world, tapping into the resources available in diverse national systems of innovation (Cantwell and Mudambi, 2005). These MNEs have created pipelines for the flow of knowledge and resources within their global internal networks (Lorenzen and Mudambi, 2013). However, they rapidly recognized that transferring knowledge out of its local context (Szulanski, 1996) and integrating it across geographic space is an enormously difficult task.
In this process, MNEs are increasingly using their subsidiaries to integrate knowledge from headquarters with that sourced from their local environment (Song and Shin, 2008) and the mobility of high skill employees as a means of enhancing the intra-firm transfer and integration of knowledge (Almeida and Kogut, 1999; Song et al., 2003).

The large and growing number of highly skilled migrants in OECD countries provides MNEs natural stock of boundary spanners (Schotter et al., 2017). For instance, major tech companies like Amazon, Microsoft and Intel are among the top recruiters of foreign-origin skilled workers in science and engineering in the US (NFAP, 2018). These highly skilled workers are embedded in the corporate network of the MNE, but almost always retain knowledge and connections in their countries of origin (CoO) (Chowdhury, 2016). Further, many of these highly skilled employees of the MNE have an interest in returning to work in their CoOs. The phenomenon of “returnee” highly skilled employees therefore raises important questions yet another potential positive impact of migrants on the economies of their destination firms. This involves studying how they can serve as bridges to their CoOs and improve the functioning of the MNE’s global innovation network.

**Theoretical framework**

R&D activities by MNEs have over time become more geographically distributed (Iversen et al., 2016; Scalera et al., 2017). While, originally, R&D used to be conducted abroad only to adapt home-developed technology to foreign markets, it has gradually gained a more creative role in MNEs’ technology strategy (Cantwell and Mudambi, 2005).

Research dating back to the seminal works by Dunning (1956) and Buckley and Casson (1976) suggests that the very existence of MNEs can be explained in light of their unique ability to transfer, integrate and exploit knowledge in the geographically distributed intra-firm network.
Unfortunately, the fact that MNEs are superior in accomplishing these knowledge-related tasks compared to external market mechanisms does not imply that intra-MNE knowledge transfer and exploitation occur effectively and with no effort (Gupta and Govindarajan, 2000). Consider the case of an MNE that is willing to source specialized knowledge that has been developed in a foreign country, to ultimately allow the home-based R&D office to integrate and re-use such knowledge for new technology creation. To source this knowledge (stage 1), the MNE can extend its R&D activities into this foreign country (hereafter referred to as the host-country), for instance through the establishment of a local R&D lab or the development of knowledge-based collaborations with local scientists and engineers. Once the knowledge has been sourced locally (i.e., in the host-country), the MNE must find ways to bring this knowledge to the home-country, so that it can be integrated and re-used by inventors working in the MNE’s home-based R&D office (stage 2). However, the integration of foreign knowledge in the MNE’s home-based R&D office, could be hindered by substantial barriers. First, the knowledge that the MNE wishes to assimilate from a host-country could be locked into the cultural, religious and linguistic context of this location (Bartholomew, 1997; Choudhury and Kim, 2017). While this might not prevent knowledge sourcing (stage 1) since this stage occurs via R&D activities developed directly in the host-country (Singh, 2008), it is likely to significantly affect knowledge integration (stage 2). Knowledge is context-specific and it accumulates in coevolution with aspects of its local environment (Nelson and Winter, 1982). Thus, knowledge developed in foreign countries is the outcome of processes that have occurred in a context that is different from the MNE’s home-country (Kogut, 1991; DiMaggio and Powell, 1991). As suggested by Teece et al. (1997), learning and assimilation become very complex when several parameters of the learning environment change simultaneously. Under this condition, the recipient’s ability to
understand cause-effect relationships is hindered, due to a lack of supporting cognitive structures (Teece et al., 1997). Thus, while MNEs may succeed in sourcing knowledge developed in a different country context because their local R&D activities permit personal interaction and exposure to the host-country context (stage 1), the subsequent stage of integration of such knowledge (stage 2) could be hindered by the limited understanding of the idiosyncratic characteristics of the geographical context and of the underlying causal relationships.

In addition to increased casual ambiguity, knowledge that is strongly embedded in the location where it has been developed is also likely to have a significant tacit component that raises its stickiness (Szulanski, 1996; von Hippel, 1994): while the codified component of knowledge moves easily, the tacit component associated with the underlying know-how can only be transferred via personal interaction and direct experience (Dasgupta and David, 1994; Kerr, 2008). While these channels can be exploited by the MNE’s inventors working in the host-country, thus favoring the knowledge sourcing process (stage 1), they are not available to the MNE’s home-based scientists (stage 2) due to the geographical distance.

Finally, the MNE’s home-based R&D office could also lack sufficient absorptive capacity and motivational disposition to assimilate and integrate foreign knowledge (Levinthal and March, 1993; Gupta and Govindarajan, 2000). Previous research has demonstrated that the so-called “not-invented-here” (NIH) syndrome (Katz and Allen, 1982) is very pervasive in MNEs. In fact, units based in a particular country might be reluctant to accept knowledge inflows from peer units based in other countries as they lack the absorptive capacity required to understand and appraise the value of this knowledge (Gupta and Govindarajan, 2000). Obstructing knowledge inflows from sister units could also be the outcome of a strategic decision. The MNE internal network can be conceived as a political coalition (Holm and Pedersen, 2000) in which subsidiaries compete for
resource allocation and thus seek to gain visibility and bargaining power. Since intra-MNE knowledge flows are a key determinant of MNE units’ relative power (Mudambi and Navarra, 2004), managers at the MNE home-based R&D office could refuse any information inflows that might insinuate that they are less competent than other units in the MNE organization, as is typical when ego-defense mechanisms (Allport, 1937; Sherif and Cantrill, 1947) perturbate the organizational environment. Power struggles (Pfeffer, 1981) within the MNE organization could also induce managers at the MNE home-based R&D office to pretend that the knowledge accessed in the host-country is not valuable or unique to discredit competing units and lessen their potential importance (Gupta and Govindarajan, 2000).

IB literature has documented that also the cross-border exploitation of the MNE knowledge does not happen automatically. To illustrate this phenomenon, consider the previous example and suppose that, once the MNE has managed to source and integrate foreign knowledge into new valuable technology, it chooses to exploit this asset abroad to better appropriate the related economic rents. The MNE’s newly created technology could be incorporated into products, processes, functions and procedures (Dunning and Lundan, 2009) that may have a value in the host-country and other foreign countries. However, as previous IB research has widely documented, the exploitation of the MNE’s technological assets in a foreign market requires significant adaptation to local specifications, manufacturing conditions, customer needs and technical requirements, and to the general competitive context (Blanc and Sierra, 1999). Even when the technology is merely supplied to the MNE’s foreign production facilities, it still requires to be tailored to country-specific regulations, codes and standards that might be in use in the foreign country (Von Zedtwitz et al., 2004). In other words, to be effective, adaptation processes should be inspired by a profound knowledge and understanding of the foreign market and general
business environment (Zaheer, 1995). However, knowledge of a given market can only be gained by means of embeddedness in the specific market environment, and exposure to the broader societal characteristics which often remain obscure to actors that are distant from the target market (Meyer et al., 2011). A lack of knowledge about foreign markets, and operations in these markets, hinders the MNE’s ability to carry out successful technology adaptation.

_The role of ethnic inventors in MNE’s knowledge integration and exploitation_

Previous IB literature suggests that in order to facilitate both knowledge integration and knowledge exploitation, the existence of ties channeling relevant knowledge that is embedded in the host-country back to the MNE home-country could be decisive (Gupta and Govindarajan, 2000). We contend that MNEs’ ethnic inventors working in cross-border R&D teams can provide these ties. Specifically, we argue that when the ethnicity of these inventors matches the country where the other members of the inventor team are located (i.e. when the ethnic inventor’s CoO matches the _host-country_ of the MNE’s R&D activities), the resulting ethnic ties may alleviate barriers to knowledge integration and exploitation, acting as channels that facilitate the flow of knowledge between the MNE home- and host-country.

Migration research has suggested that migrants and their descendants remain involved in familial, social, religious, economic, political and cultural processes that extend across distances and borders, even when they settle down in a given country of destination (CoD) (Basch et al., 1994; Faist, 2000; Portes et al., 1999). These enduring ties that connect them to their CoO are preserved while migrants embed in their CoD, which renders their social capital transnational (Nowicka, 2014). Transnational social capital helps ethnic inventors to retain high levels of absorptive capacity of the knowledge developed in their CoO (Zahra and George, 2002), while simultaneously being able to integrate and recombine such knowledge with technological inputs originating in the
MNE’s home-country. In fact, ethnic migrants have a significant understanding of both the institutional contexts involved in the dynamics of knowledge integration (i.e., the host-country or CoO) and the context in which knowledge should be transferred and integrated (i.e., the home-country or CoD). Because they are able to recognize and appreciate the value of the knowledge developed in their CoO, they are likely to be more open to and collaborative with members of the inventor team who are located in their CoO, thus being immune to the NIH syndrome (Katz and Allen, 1982). The ethnic ties that link own-ethnic inventors to the CoO may also increase their motivational disposition to integrate knowledge originating from the CoO. In fact, own-ethnic inventors have no incentives to picture the knowledge originating from their CoO as poor and valueless, since this would also reduce the perception of their own competences in the MNE network. Due to cultural, linguistic, religious and institutional proximity with the other inventors, their presence in the team tends to facilitate trust, coordination and, in turn, learning (Boschma, 2005), which are fundamental prerequisite for effective integration of existing technological inputs into new knowledge. The existence of ethnic ties creates a common background and facilitates the emergence of a shared working culture that are likely to reduce the amount of resources (e.g., time) needed to develop effective interaction and communication routines. For instance, ethnic ties may be expected to reduce the likelihood of misunderstandings across team members and, in turn, increase the efficiency of the team’s collaborative dynamics. Furthermore, the presence of an ethnic inventor in core positions in the MNE network, as is when the ethnic inventor works in the home-office, is likely to provide more legitimacy to the consistent minority represented by the foreign inventors, thus empowering its participation, favoring creativity and idea sharing, reducing the negative effects of social categorization and smoothing the social processes, so to ultimately increase its contribution to the knowledge creation processes (Pfeffer, 1983; Tajfel, 1982).
Based on this reasoning, we expect that the presence in the inventor team of an own-ethnic inventor from a certain CoO favors the transfer and integration of foreign technological knowledge and, more specifically, of technological knowledge from that CoO. In fact, due to their enduring social ties to the CoO, greater absorptive capacity and motivational disposition to source knowledge from their CoO, own-ethnic inventors can help unlock foreign technological knowledge that was previously embedded within the institutional contexts of their CoO, and recombine it with technological inputs available within the MNE to ultimately develop new knowledge.

For similar reasons, ethnic inventors should also be better equipped to understand the specificities of the CoO market. Ethnic inventors enjoy greater exposure to and awareness of the CoO’s habits, cultural and industry norms, manufacturing conditions and institutional context (D’ambrosio et al., 2018), thereby being well positioned to understand and interpret the needs of the ethnic demand and the rules and specificities that govern the CoO general business environment. As a consequence, they may provide critical inputs to processes of technology adaptation. More specifically, migration research highlights that migrant workers actively seek to validate their “cultural capital” in their CoD (Erel, 2010; Nowicka, 2014), thereby using their “ethnic belonging” via social and inter-ethnic networks as a resource in their job environment (Nohl et al., 2010). In other words, migrant workers often leverage their ethnic origin as a valuable asset to improve their own status and role in the receiving organization. This suggests that ethnic inventors might be willing to exploit the opportunity to leverage their ethnic origin to acquire unique roles as facilitators of those processes through which knowledge of the CoO’s habits, preferences, industry norms and local specifications is transferred to the MNE headquarters. This should

---

4 It is worth noting that migrants who have moved across different countries of destination are likely to have knowledge and understanding of multiple country contexts. This broadens the transnational reach of their social capital (Nowicka, 2014), renders more open, flexible capable to manage cultural shocks (Franzoni et al., 2014) and, ultimately, puts them in a privileged position to serve as channels for knowledge sourcing originating from different geographies.
facilitate the adaptation and exploitation of the newly-created technology in foreign markets and, more specifically, in the CoO market.

**Empirical strategy**

*Data and methods*

The objective of this paper is twofold. First, we aim at analyzing the role of ethnic inventors in fostering the integration of foreign knowledge in the intra-organizational environment. Specifically, we test whether the existence of a specific typology of tie between the MNE’s US-based ethnic inventor and inventors based in the MNE’s host-country, i.e. the membership in the same ethnic community⁵, ameliorates the transfer and subsequent re-use of knowledge embedded in the host-country⁶ into the MNE’s headquarters. Second, we explore whether the integration of host-country-specific knowledge by means of an (own)ethnic inventor is also linked to the exploitation of the resulting innovation in that country market.

In order to achieve our research objectives we use USPTO patent data granted to US-based MNEs operating in broad knowledge-intensive sectors (computers and communication, electrical and electronics, drugs and medical, and chemical) over the period 1975-2009.

To implement our theoretical model, we follow the leading approach of using backward citations as a “noisy” proxy of knowledge flows (Jaffe et al., 1993; Almeida, 1996). In the USPTO framework, patent applicants are legally required to fully list the relevant prior art to delineate the conferred property rights. Thus, citations make possible to track knowledge that preceded the

---

⁵ According to Breschi et al. (2017), when defining ethnic ties between inventors we explicitly refer to (social) ties whose existence is independent of the organizational, technological, personal inventors’ background and/or their co-location experience.

⁶ We use the terms “CoO” and “host-country” interchangeably to refer to the country of origin of the ethnic inventor which, in the case of own-ethnic inventors, overlaps with the host-country, i.e. the foreign country in which the MNE carries out local R&D activities in order to source locally embedded knowledge.
moment of the invention (OECD, 2009) and has been used as an input to the new technology creation process. While useful for exploring knowledge flows, patent citations might contain connections of spurious nature. In particular, there might be the case in which knowledge flows occur without generating a citation or citations do not imply a knowledge flow (Jaffe et al., 1993). However, previous works agree that it is reasonable to argue that backward “citations [...] are sufficiently correlated with knowledge flows to allow statistical analysis of the proxies to be informative regarding the underlying phenomenon of interest” (Jaffe and Trajtenberg, 2002: 379).

The geography of innovation literature suggests that knowledge flows are strongly tight to locations and related geographical space (Breschi, 2011; Henderson, 1997). The main empirical challenge associated to the analysis of the drivers of knowledge flows is thus to account for correlations that might come from the pre-existent spatial distribution of patenting activities and its evolution over time. Jaffe et al. (1993) have pioneered a methodology (hereafter JTH method) to tackle this issue leveraging the information included in patent citations, and using two sets of patent pairs employed as case and control sets, respectively. Further improvements to the traditional JTH method have been implemented, proposing -for example- a higher level of detail in the matching of the technological classification of patents (Thompson and Fox-Kean, 2005).

More recently, research has focused on analyzing more fine-grained mechanisms underlying the pattern of knowledge flows (e.g., Breschi and Lissoni, 2005, 2009). In fact, it has been empirically documented that citations are significantly driven by various types of social ties (e.g. Agrawal et al., 2006; Breschi and Lissoni, 2009). Within this strand literature, the role of ethnic ties between inventors is gaining growing attention (Breschi et al., 2017; Almeida et al., 2015; Agrawal et al., 2008, 2011).
Following the recent approaches, our methodology lies at the intersection between studies in geography of innovation and research on migration, by adopting an extended JTH method used also by Agrawal et al. (2011). More specifically, we adapt the JTH method to match every backward citation to a control patent that perfectly mimics the technological and temporal distribution of their original counterpart. Selecting a counterfactual group provides us with a benchmark that control for the pre-existing spatial distribution of innovative activity across time and technological space. In other words, if the ethnic inventors play no role in facilitating the international knowledge sourcing activities, then we would not observe any significant disproportion in the probability that a citation link refers to an actual citation or a matched observation. Otherwise, we can use the estimated coefficients as a measure of causal effects of the presence of an ethnic inventor in cross-border knowledge integration phenomena. We discuss in details the sampling and matching procedures in the next Section.

**Sampling and counterfactual build**

For our empirical analysis, we select our sample of focal patents starting from the USPTO patent records collected in the public available database designed by Kerr and Kerr (2018). This dataset contains information about industrial patents granted to US-based MNEs between January 1975 and May 2009. According to Kerr and Kerr (2018) conceptual and empirical framework, US-based MNEs included in the sample are defined as public companies conducting global technology development via “entering into patenting abroad after first patenting in the US” (p. 10). Exploiting commercial ethnic names databases and name-matching algorithms, the authors determined the probable ethnicities of the inventors listed in the patents (Kerr and Kerr, 2018; for more details about the matching procedure, see Kerr, 2007 and 2010). The procedure distinguishes
between nine ethnicities: Anglo-Saxon, Chinese, European, Hispanic, Indian, Japanese, Korean, Russian and Vietnamese. Ethnic inventors are defined as US-based inventors (according to the address of the inventors provided by the patent document), whose name can be matched to one of the previous ethnicities. For each focal patent, the ethnicity assignment equals the average of the ethnic probability assigned to each inventor in the research team. Therefore, we identify the ethnic base of the foreign-origin inventors based on the highest probability score at the patent level.

Since we aim at investigating the role of ethnic migrants in the intra-MNE knowledge flows, we focus on *global collaborative patents*\(^7\) granted to US-based MNEs operating in knowledge-intensive industries and with at least one non-Anglo-Saxon ethnic inventor in the team. The selection of MNEs in knowledge-intensive sectors is driven by some general considerations. First, the share of global collaborative patents with ethnic inventors with respect to the total US patenting activity has both higher levels and growth rates in these sectors than in the residual technological fields (Kerr and Kerr, 2018). Second, these technological categories are related to the industries featuring higher levels of technology intensity and innovation (OECD, 2011). Third, the high levels of mobility of high skilled workers in these sectors make them a suitable setting for studying the phenomenon at hand.

In order to pinpoint MNEs operating in knowledge-intensive industries we rely on the main technological category (Hall et al., 2001) of the patents included in each firm portfolio as recorded in the Kerr and Kerr (2018) dataset. Consequently, we classify a firm as belonging to a specific sector if it registers at least 20% of its overall patent stock in the related technological category during the period 1975-2009 (for a similar approach, Jiang et al., 2011).\(^8\) If a firm falls in more than

---

\(^7\) Global collaborative patents are defined as patents where at least one inventor is located within the US and at least one resides in a foreign country (Kerr and Kerr, 2018).

\(^8\) Previous works used different methods to identify the technological specialization of a firm starting from its patent portfolio (Frost and Zhou, 2005; Jiang et al., 2011). Given the distribution of the patents’ technological classification
one classification based on the previous criterion, we assign the firm to the more represented sector in the firm’s patent portfolio based on the patents’ technological category. For each patent, in order to identify the foreign country where the MNE’s innovative activities took place (i.e., host-country) we use the most frequent location of non-US inventors included in the patent inventor team, following the procedure proposed by Kerr and Kerr (2018). For the purpose of our study and the application of the JTH method, we also impose that all focal patents included in our sample cite at least another USPTO patent. Further, to build our sample we include both assignee and inventor self-citations to be able to explore the role of ethnic migrants in the intra-organizational knowledge flows, in line with the approach of Agrawal et al. (2011)\(^9\). More generally, previous studies have widely used self-citations to measure intra-firm knowledge flows (e.g., Frost and Zhou, 2005), as they allow to capture the degree to which MNEs use knowledge developed in different units of their organization. However, we also perform robustness checks by controlling for self-citations and non-self-citations as we acknowledge that their citation patterns might differ. Additionally, in order to test the potential role of ethnic inventors in MNE cross-border knowledge exploitation, we recover information on the geographic extension of patent protection using the ORBIT database, provided by QUESTEL. In particular, we collect data about the patent family to which each focal patent belongs. According to the FamPat database provided by ORBIT, a family groups together all the patent documents referring to the same single invention. Based on the European Patent Office (EPO)’s strict family rule, the ORBIT FamPat database aggregates patent

\(^9\) It is worth noting that a self-citation (especially at assignee-level) does not necessarily imply an actual flow of knowledge between the MNE’s subunits. On the one hand, knowledge may be transferred from one subunit to another via other more informal channels, such for face-to-face interactions, and therefore it does not necessarily materialize in a citation; on the other hand, a self-citation may be included even if not associated to an actual transfer or knowledge, but rather for accomplishing more strategic motivations. However, according to the argument of Frost and Zhou (2005), we expect that, other things being, two MNEs’ subunits frequently citing each other’s works are also more likely to share and use their respective innovation compared to two subunits that do not cite each other’s patents.
records from many Patent Offices across the world having exactly the same priority or combination of priorities (equivalents). Since each patent document is assigned to only one group, no single patent number may appear in two distinct families. Because of the territorial nature of patent protection, patent applicants who wish to protect their invention in foreign markets are required to file an application in each foreign country in which protection is sought (Martinez, 2010). Thus, MNEs that plan to exploit their technological assets internationally either to supply the new technology to their own foreign production facilities or to incorporate the technology into products or processes that target the foreign demand can be expected to extend the patent protection into the foreign country (Blind et al., 2006). We classify an innovation as protected in a certain host-country, if the patent family of the focal patent contains at least one publication in the filing office of the host-country or if the country appears as a designated state in an EPO, World Intellectual Property Organization (WIPO) or Patent Cooperation Treaty (PCT) patent document. Since the economic and institutional setting of the host-countries in our sample is highly heterogeneous, we also distinguish whether the host-country where the MNE sought protection is an OECD member state or not.

To implement the matching procedure described in the Section 3.1, for each patent cited by our focal patent we randomly identify a control patent among the USPTO granted patents that satisfies the following criteria: 1) it has the same application year as the backward citation; 2) it reports exactly the same USPTO technological classification (same number of classes and respective codes) as the backward citation; 3) it is not a reference of the focal patent. If we are not able to find any perfect match, we exclude the focal-citation observation from the sample. Our level of analysis is the pair of patent-citation and the described procedure leaves us with 316,050

---

10 For more detailed definitions of patent families, see Martinez (2010).
observations (158,025 patent-citation pairs and 158,025 patent-control pairs). The focal sample contains 11,051 patents, of which 3,016 report an own-ethnic migrant in their list of inventors. On average, each focal patent cites 91.65 documents, but the statistic increases to 142.10 if we consider the group of global collaborative patents with an own-ethnic migrant in the inventor team.

Regression model

We run the following regression, as our baseline model, by means of a Linear Probability Model (LPM) with standard errors clustered at the focal patent level\(^{11}\):

\[
P(Citation = 1) = \alpha_0 + \alpha_1 Foreign + \alpha_2 HostCountry + \alpha_3 OwnEthnic + \delta_i + \tau_t + \eta_j + \epsilon_{ijt},
\]

where the observations are backward citations pairs and our dependent variable Citation is dichotomous taking the value of 1 if the citation represents an actual citation included in the focal patent, i.e. respecting a knowledge flow, and 0 if it is a matched observation. As for the regressors, we include the following main dummy variables: Foreign equals to 1 if at least one inventor in the citation/control is based outside the US and 0 otherwise; HostCountry takes value of 1 if at least one inventor in the citation/control is based in the MNE’s host-country, , i.e., representing a co-location tie, and 0 otherwise; OwnEthnic is equal to 1 if the ethnic inventor in the focal patent has the same ethnic base as the MNE’s host-country, i.e., representing an ethnic tie, and 0 otherwise.

Table 1 reports the descriptive statistics for the regression sample.

\(^{11}\) According to other works which used a similar approach (e.g., Breschi et al., 2017), we privileged the use of LPM over probit/logit models, as the former provide a more direct interpretation of the estimated coefficients, which directly represent the marginal effects. Logit estimates, which provide similar results, are available upon request.
Throughout our analysis, we also introduce model specifications that include interaction terms between OwnEthnic and the other main regressors to test for any moderation effect of the existence of an ethnic tie between the foreign-origin inventor and the host-country where the R&D activities of the MNE take place. In other words, we analyze whether the presence of an own-ethnic inventor in the focal patent relates with more intense integration of cross-border knowledge, and specifically with more knowledge flows with the host-country, than inventor teams with a generic ethnic migrant. All the regressions contain firm $\delta_i$, application year $\tau_t$ and main technology $\eta_j$ fixed effects.

As robustness checks, we also 1) run a model that includes host-country fixed effects or ethnic base of the host-country fixed effects; 2) control for both assignee and inventor self-citations; 3) perform a subsample analysis to test for heterogeneity in the assignee self-citation subgroups.

Finally, in order to accomplish our second research objective and test for any potential heterogeneity in the role of ethnic inventors in terms of exploitation of innovation in foreign countries, and more particularly in the CoO, we conduct a subsample analysis in which we split our sample according to whether the global collaborative patents are protected, and thus exploited, in the host-country/CoO. To control for potential heterogeneity arising from the characteristics of the host-country context, we further split the group of patents protected in the host-country by separating countries belonging to the OECD or not. In doing so, we seek to account for cross-country economic, social and institutional characteristics that might influence the decision of the MNEs to protect the innovation in the host-country.

Results
Table 3 reports the main results for the full sample, which are primarily used to test our argument that ethnic inventors facilitate the integration of foreign knowledge and, more specifically, of knowledge developed in the host-country (or the ethnic inventor’s CoO). Focusing on Model 1, we find evidence that global collaborative patents with ethnic migrants are less likely to integrate cross-border knowledge than the control group (as captured by backward citations to patents developed outside of the US), as suggested by the negative and significant coefficient of the variable \textit{Foreign}. Thus, the presence of an ethnic inventor in the inventor team does not \textit{per se} increase the MNE ability to pursue international knowledge sourcing strategies, while on the other end it seems that –in general- local knowledge flows are more frequent. However, the large, positive and significant coefficient of \textit{HostCountry} also indicates that global collaborative patents with ethnic migrants integrate more host-country-specific knowledge (as captured by backward citations to patents developed in the host-country). Moreover, the magnitude of the effect related to the integration of knowledge with a specific host-country base is greater than the integration of generic foreign knowledge as indicated by the size of the coefficients. Taken together, this evidence seems to point toward a pattern of international knowledge sourcing that emphasizes location-specific knowledge and, more specifically, host-country-specific knowledge, rather than involving a more geographically diversified set of knowledge inputs. While, from a theoretical viewpoint, the characteristics of the ethnic migrants might suggest that their open-mindedness, their cultural rich background and their transnational capital could render the ethnic migrants more versatile so to generally reinforce MNEs’ ability to recombine geographically distributed knowledge inputs, it seems that the specificity of their ethnic ties is the predominant effect.

The results are further confirmed when controlling for the inventor team of the focal patent including an ethnic inventor with a specific ethnic tie with the host-country (Model 2). The variable
*OwnEthnic* positively affects the likelihood of observing the occurrence of a knowledge flow. In other words, global collaborative patents with own-ethnic migrants have a greater number of backward citations than global collaborative patents involving ethnic migrants without a specific ethnic tie with the host-country.

Introducing the interaction effects of *OwnEthnic* with both *Foreign* and *HostCountry* further supports the theoretical interpretation of our results. In fact, Model 3 confirms that global collaborative patents with at least one own-ethnic inventor are associated with more intense cross-national knowledge sourcing. In Model 4, the positive and significant interaction coefficient also suggests that these patents specifically integrate more knowledge emanating from the CoO in the intra-corporate context. The combined interpretation of these specifications seems to confirm the own-ethnic inventor’s role as reinforcing channel for international knowledge sourcing focused on a specific geography. More specifically, these results provide support for the key role own-ethnic inventors play as enablers of the integration of host-country knowledge in the intra-corporate network. The visual representation of the average marginal effects of the main regressors at different levels of *OwnEthnic* corroborates this assumption (Fig. 1).

The introduction of host-country or ethnic base fixed effects (Table 4) does not significantly alter the estimated coefficients with respect to the model containing the interaction between *OwnEthnic* and *Foreign* (Models 2 and 5). Even though the *OwnEthnic* coefficients in Models 3 and 6 change their sign and they are now negative, the presence of positive and significant interaction coefficients with a higher magnitude than *OwnEthnic* also confirms the reinforcing role of an ethnic tie in knowledge integration in these specifications. In both Models the linear combination of the coefficients of the variables *OwnEthnic, HostCountry* and *HostCountry* *OwnEthnic* turns out to be positive and significant, reassuring about the robustness of the results. The interpretation
is confirmed by the average marginal effects of *HostCountry* and *Foreign* computed at the different values of the variable *OwnEthnic*, whose graphical representation in Figure 2 depicts a positive moderator effect of *OwnEthnic* for both the regressors. Overall, these results confirm the role of the own-ethnic inventor, even if clearly heterogeneity exists due to the specific characteristics of the host-countries and their institutional, cultural and technological frameworks.

In Table 5, we introduce controls for the instances where the focal patent cites a patent that is assigned to the same assignee (assignee self-citations) or where the focal patent cites a patent that shares at least one inventor in the inventor team (inventor self-citations). As discussed in the previous Section, the decision to include in the final sample self-citations (deriving from both assignee and/or inventor) enables us to capture intra-firm knowledge flows, which represent an important part of the knowledge sourcing strategies of MNEs. In their knowledge creation processes, MNEs, in fact, tend to include not only external foreign knowledge, but also knowledge developed within their organizational boundaries by host-country subsidiaries distributed in foreign locations. As a result self-citations allow to measure for headquarter-subsidiary or subsidiary-subsidiary knowledge flows, which represent a way through which MNEs internalize the knowledge output resulting from their geographically distributed R&D efforts. However, we acknowledge that the citation patterns of self-citations might differ from non-self-citations, and the analysis in Table 5 has a twofold objective, i.e. to show the robustness of our main results when controlling for the self-citation effect and to discuss more in detail the role of ethnic migrant in intra-firm knowledge flows. While patterns and significance of the coefficients are generally confirmed, the *OwnEthnic* coefficient loses significance in the specification which includes its interaction with *HostCountry* (Model 3 and 6). Furthermore, the limited changes in the magnitude
of the coefficients compared to the baseline results also suggest that there exists a correlation between the main regressors and the citations being an assignee or inventor self-citation. Therefore, we explicitly test for any heterogeneity in the phenomenon due to the nature of the citations, and we perform a subsample analysis separating the dyads over the assignee self-citation dimension. We specifically focus only on assignee self-citations (rather than also inventor self-citations), as we aim at testing whether a specific intra-firm knowledge transfer mechanism exists. In Table 6, Models 1-3 show the results from a LPM regression applied to the sample containing pairs corresponding to non-self-citations, while Models 4-6 refer to the group of self-citations. The pattern and significance of the coefficients are generally confirmed, but interestingly enough the \( Host\text{Country}*Own\text{Ethnic} \) interaction term is positive and strongly significant only in the assignee self-citation subsample. In this latter specification, the sign of \( Own\text{Ethnic} \) also changes with respect to the baseline model, but the average marginal effect at different levels of \( Own\text{Ethnic} \) is confirmed to be higher for the patents with an ethnic tie with the host-country (Fig. 3). Furthermore, the magnitude of all the coefficients is greater in this specification than in the group excluding all the assignee self-citations. Thus, these results demonstrate that the overall knowledge integration effect associated to own-ethnic inventors is bounded to the flows of host-country knowledge developed within the MNEs’ organizational boundaries. In other words, the role of the own-ethnic inventors in host-country knowledge integration is mainly activated via an intra-organizational conduit leveraging the ethnic ties to assimilate knowledge developed in the host-country but within the MNE network.

Finally, we investigate the role of own-ethnic inventors in facilitating the exploitation of new technology resulting from the integration of host-country knowledge in the CoO market. In Table 6, we split our main sample based on whether the MNEs choose to seek protection for the global
collaborative patents only in the MNE home-country, i.e. the US (Models 1-3), or also in the host-country (Models 4-6). According to our theoretical arguments, we should observe that, in the second subsample, knowledge integration from the host-country through the channel of own-ethnic inventors is greater. The existence of this pattern would depict an association between the intensity of knowledge integration through ethnic inventors and the MNE’s ability to gain access to information on the host-country market, industry norms and general business environment. Specifically, it would signal a more effective process of adaptation ultimately leading to exploitation of the resulting technology in the host-country, as highlighted by an extension of patent protection. However, the empirical analysis does not provide support for such role of own-ethnic inventors. In fact, both the OwnEthnic coefficients in Models 5 and 6 and their interactions with the other main regressors are not statistically significant in the subsample of the inventions protected in the host-country. In order to better investigate this result, in Table 7 we further split the group of patents protected in the host-country distinguishing countries being OECD members (Models 7-9) or not (Models 4-6). This procedure allows to account for cross-country economic, social and institutional heterogeneity of the host-countries, as some location-specific characteristics might influence the decision of the MNEs to exploit innovation in a certain foreign market. For instance, it is possible that while the MNE is willing to exploit the innovation in the CoO, it chooses not to extend its formal patent protection in that country simply to avoid being exposed to knowledge expropriation risks arising from a low degree of IPR protection in the country. Once again, the results do not provide any empirical support for the potential “knowledge exploitation” role of own-ethnic inventors. Moreover, the HostCountry coefficients (and thus the related knowledge integration phenomena) seem to be stronger in the case where the underlying innovation lacks IPR protection in the host-country (Models 4-6). Collectively, these final results
confirm the lack of a significant moderation effect related to the role of the own-ethnic inventor in MNEs’ knowledge exploitation, regardless of the institutional profile of the country involved in the knowledge creation process. Thus, while recent literature suggests that migrants open up new markets for CoD companies, our analysis seems to suggest that own-ethnic inventors do not serve this function; they seem to act as effective carriers of technological knowledge, but not of market- and business-related information (Cowan et al., 2000).

Conclusions
This study contributes to the nascent stream of literature on the contribution of highly skilled migrants to the knowledge creation processes of MNEs (Choudhury, 2016). Answering to the recent call for greater consideration of firms in studies on migration (Kerr et al., 2015), we show that the presence of an own-ethnic inventor in cross-border R&D teams (captured by global collaborative patents) triggers a disproportionate use of knowledge developed in the host-country within the intra-MNE network. This finding seems to suggest that own-ethnic inventors serve a bridging function between the host- and the home-country of the MNE.

Ethnic ties help to mobilize the knowledge inputs that MNEs source abroad via local R&D activities and that could be difficult to assimilate by inventors located in the MNE R&D home-office, due to barriers to knowledge transfer across borders and organizational frictions. Thanks to their enduring ties to the CoO, ethnic inventors enjoy greater exposure to and understanding of the context in which foreign knowledge has been developed; in turn, this provides them with absorptive capacity and motivational disposition to source knowledge from the CoO. As a consequence, ethnic inventors facilitate assimilation of knowledge that, albeit successfully sourced in the host-country, could face significant obstacles in its journey to the MNE home-based R&D
office. Thus, it could be argued that ethnic ties established through ethnic inventors complement knowledge sourcing channels based on MNE co-location with the foreign knowledge source, and allow to fulfill the knowledge integration process within the MNE network.

Additionally, our empirical analysis also clarifies that ethnic inventors’ bridging function is limited to knowledge developed within the MNE. When sourcing knowledge developed outside of the firm organizational boundaries, even own-ethnic inventors cease to act as effective channels of cross-border knowledge integration. This confirms the superior role of firms compared to external market mechanisms in the transfer of knowledge across borders, and suggest that ethnic ties – *per se* – might not be sufficiently powerful to overcome the barriers that are inherent to the cross-national movement of technology.

On the other hand, the empirical analysis does not support the knowledge exploitation role of own-ethnic inventors. In fact, we find no significant evidence of an association between the intensity of knowledge integration through ethnic inventors and the likelihood of exploitation of the resulting new technology in the CoO. This seems to suggest that own-ethnic inventors are effective carriers of technological *knowledge*, but not of *information* about the foreign market, its operations and the general business environment (Cowan et al., 2000). Hence, this finding provides some preliminary insights into the recent debate on the *content* of diffusion flows activated by highly skilled migrants (Lissoni, 2018).

We believe our findings have relevant implications for the management of high-skilled workers in MNEs. If ethnic ties provide effective channels to ensure the integration of knowledge developed in their CoO within the intra-corporate network, MNEs could strategically allocate ethnic inventors to R&D teams that involve inventors based in the CoO when they need to use specialized technology that is embedded in such geographical context. More generally, our study also speaks
to the debate on the influence of high-skilled migration on a country’s domestic workforce. Specifically, it shows that due to their unique endowment with enduring social ties to their CoO and persistent exposure to the CoO’s institutional and knowledge environment, foreign-origin scientists serve some functions within the MNE that domestic scientists might not be equally able to carry out. In other words, foreign-origin skilled migrants are unlikely to crowd-out a receiving country’s domestic workforce, simply because the two groups of talents are not perfectly substitutable. Policy makers could also use this insight to support MNEs’ recruiting of advanced foreign human capital in key industries, by defining targeted policies to attract high-skilled migrants (via preferential access to visa, fiscal incentives, etc) from specific regions that match or complement their countries’ technological specialization or that are home to strategic clusters or global centers of technical excellence.

Our study also has limitations. For instance, our empirical analysis does not allow to highlight potential variation in the contribution of ethnic inventors to MNEs’ knowledge creation processes across first and subsequent generations of migrants. Moreover, we do not have information on communication and knowledge sharing systems that different MNEs included in our sample might have adopted, which could reveal the existence of firm-specific arrangements that may further reinforce or hamper the bridging function of ethnic inventors. Future research should continue to explore these issues and shed more light on the role firms, and particularly MNEs, as key actors that determine and direct highly skilled migration flows.
References


Tables and Figures

Table 1 – Classification of migration flows according to origin and destination countries.

<table>
<thead>
<tr>
<th>Origin country</th>
<th>Destination country</th>
<th>OECD</th>
<th>Poor / emerging</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD</td>
<td>1. Mainly high skill migrant flows</td>
<td></td>
<td>4. Negligible</td>
</tr>
<tr>
<td>Poor / emerging</td>
<td>2. Mix of high skill and low skill</td>
<td></td>
<td>3. Mainly fleeing persecution</td>
</tr>
</tbody>
</table>

Table 2 - Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign</td>
<td>316050</td>
<td>0.37</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HostCountry</td>
<td>316050</td>
<td>0.06</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OwnEthnic</td>
<td>316050</td>
<td>0.31</td>
<td>0.46</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Grant year - focal patent</td>
<td>316050</td>
<td>2003.30</td>
<td>5.19</td>
<td>1976</td>
<td>2009</td>
</tr>
<tr>
<td>Grant year - citation</td>
<td>316050</td>
<td>1994.63</td>
<td>8.08</td>
<td>1975</td>
<td>2010</td>
</tr>
</tbody>
</table>
Table 3. The role of (own-)ethnic migrants in knowledge sourcing - LPM regression

<table>
<thead>
<tr>
<th></th>
<th>Dep. Var. = Citation</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign</td>
<td>-0.1350***</td>
<td>-0.1350***</td>
<td>-0.1487***</td>
<td>-0.1350***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0030)</td>
<td>(0.0032)</td>
<td>(0.0030)</td>
<td></td>
</tr>
<tr>
<td>HostCountry</td>
<td>0.2207***</td>
<td>0.2208***</td>
<td>0.2215***</td>
<td>0.2117***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0054)</td>
<td>(0.0054)</td>
<td>(0.0054)</td>
<td>(0.0068)</td>
<td></td>
</tr>
<tr>
<td>OwnEthnic</td>
<td>0.0033***</td>
<td>-0.0124***</td>
<td>0.0016*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0022)</td>
<td>(0.0008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign*OwnEthnic</td>
<td>0.0434***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HostCountry*OwnEthnic</td>
<td></td>
<td>0.0321**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0101)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.5757***</td>
<td>0.5730***</td>
<td>0.5679***</td>
<td>0.5745***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0098)</td>
<td>(0.0099)</td>
<td>(0.0104)</td>
<td>(0.0099)</td>
<td></td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Technology FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0185</td>
<td>0.0185</td>
<td>0.0189</td>
<td>0.0186</td>
<td></td>
</tr>
<tr>
<td>Focal Patents</td>
<td>11056</td>
<td>11056</td>
<td>11056</td>
<td>11056</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>316050</td>
<td>316050</td>
<td>316050</td>
<td>316050</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors clustered at the focal patent level in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001
<table>
<thead>
<tr>
<th>Dep. Var. = Citation</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign</td>
<td>-0.1363***</td>
<td>-0.1501***</td>
<td>-0.1363***</td>
<td>-0.1360***</td>
<td>-0.1495***</td>
<td>-0.1360***</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0032)</td>
<td>(0.0030)</td>
<td>(0.0030)</td>
<td>(0.0032)</td>
<td>(0.0030)</td>
</tr>
<tr>
<td>HostCountry</td>
<td>0.2288***</td>
<td>0.2294***</td>
<td>0.2209***</td>
<td>0.2274***</td>
<td>0.2281***</td>
<td>0.2196***</td>
</tr>
<tr>
<td></td>
<td>(0.0055)</td>
<td>(0.0055)</td>
<td>(0.0069)</td>
<td>(0.0054)</td>
<td>(0.0054)</td>
<td>(0.0068)</td>
</tr>
<tr>
<td>OwnEthnic</td>
<td>-0.0015*</td>
<td>-0.0175***</td>
<td>-0.0029***</td>
<td>-0.0017*</td>
<td>-0.0176***</td>
<td>-0.0032***</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0022)</td>
<td>(0.0008)</td>
<td>(0.0007)</td>
<td>(0.0002)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Foreign*OwnEthnic</td>
<td>0.0435***</td>
<td>0.0431***</td>
<td>0.0435***</td>
<td>0.0431***</td>
<td>0.0431***</td>
<td>0.0431***</td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td>(0.0060)</td>
<td>(0.0060)</td>
<td>(0.0060)</td>
<td>(0.0060)</td>
<td>(0.0060)</td>
</tr>
<tr>
<td>HostCountry*OwnEthnic</td>
<td>0.0270**</td>
<td>0.0270**</td>
<td>0.0270**</td>
<td>0.0271**</td>
<td>0.0271**</td>
<td>0.0271**</td>
</tr>
<tr>
<td></td>
<td>(0.0100)</td>
<td>(0.0100)</td>
<td>(0.0100)</td>
<td>(0.0100)</td>
<td>(0.0100)</td>
<td>(0.0100)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.6024***</td>
<td>0.5978***</td>
<td>0.6035***</td>
<td>0.5753***</td>
<td>0.5701***</td>
<td>0.5766***</td>
</tr>
<tr>
<td></td>
<td>(0.0109)</td>
<td>(0.0115)</td>
<td>(0.0109)</td>
<td>(0.0099)</td>
<td>(0.0104)</td>
<td>(0.0099)</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Technology FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ethnic base FE</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Host country FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0189</td>
<td>0.0192</td>
<td>0.0189</td>
<td>0.0188</td>
<td>0.0192</td>
<td>0.0188</td>
</tr>
<tr>
<td>Focal Patents</td>
<td>11056</td>
<td>11056</td>
<td>11056</td>
<td>11056</td>
<td>11056</td>
<td>11056</td>
</tr>
<tr>
<td>Observations</td>
<td>316050</td>
<td>316050</td>
<td>316050</td>
<td>316050</td>
<td>316050</td>
<td>316050</td>
</tr>
</tbody>
</table>

Standard errors clustered at the focal patent level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Table 5. Robustness check II: controlling for assignee and inventor self-citations – LPM regression

<table>
<thead>
<tr>
<th></th>
<th>Assignee self-citations</th>
<th>Inventor self-citation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Foreign</td>
<td>-0.1369***</td>
<td>-0.1509***</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0032)</td>
</tr>
<tr>
<td>HostCountry</td>
<td>0.2277***</td>
<td>0.2285***</td>
</tr>
<tr>
<td></td>
<td>(0.0056)</td>
<td>(0.0056)</td>
</tr>
<tr>
<td>OwnEthnic</td>
<td>0.0031***</td>
<td>-0.0129***</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0022)</td>
</tr>
<tr>
<td>Foreign*OwnEthnic</td>
<td>0.0441***</td>
<td>0.0438***</td>
</tr>
<tr>
<td>HostCountry*OwnEthnic</td>
<td>0.0355***</td>
<td>0.0360***</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>Self-citation</td>
<td>-0.0244***</td>
<td>-0.0248***</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.5776***</td>
<td>0.5726***</td>
</tr>
<tr>
<td></td>
<td>(0.0106)</td>
<td>(0.0112)</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Technology FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.0189</td>
<td>0.0193</td>
</tr>
<tr>
<td>Focal Patents</td>
<td>11056</td>
<td>11056</td>
</tr>
<tr>
<td>Observations</td>
<td>316050</td>
<td>316050</td>
</tr>
</tbody>
</table>

Standard errors clustered at the focal patent level in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001
Table 6. Subsample analysis of assignee self-citations – LPM regression

<table>
<thead>
<tr>
<th></th>
<th>Non self-citations</th>
<th>Self-citations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Foreign</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0844***</td>
<td>-0.0925***</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0034)</td>
</tr>
<tr>
<td>HostCountry</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0627***</td>
<td>0.0639***</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0073)</td>
</tr>
<tr>
<td>OwnEthnic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0017***</td>
<td>-0.0078***</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0020)</td>
</tr>
<tr>
<td>Foreign*OwnEthnic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0253***</td>
<td>-0.0228</td>
</tr>
<tr>
<td></td>
<td>(0.0055)</td>
<td>(0.0161)</td>
</tr>
<tr>
<td>HostCountry*OwnEthnic</td>
<td>-0.0287</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5506***</td>
<td>0.5474***</td>
</tr>
<tr>
<td></td>
<td>(0.0075)</td>
<td>(0.0081)</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Technology FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Standard errors clustered at the focal patent level in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001
Table 7. Knowledge sourcing and protection in the host country - LPM regression

<table>
<thead>
<tr>
<th>Dep. Var. = Citation</th>
<th>Non-protected in host country</th>
<th>Protected also in host-country</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Foreign</td>
<td>-0.1334***</td>
<td>-0.1492***</td>
</tr>
<tr>
<td></td>
<td>(0.0033)</td>
<td>(0.0035)</td>
</tr>
<tr>
<td>HostCountry</td>
<td>0.2417***</td>
<td>0.2422***</td>
</tr>
<tr>
<td></td>
<td>(0.0059)</td>
<td>(0.0060)</td>
</tr>
<tr>
<td>OwnEthnic</td>
<td>0.0030***</td>
<td>-0.0142***</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0023)</td>
</tr>
<tr>
<td>Foreign*OwnEthnic</td>
<td>0.0478***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0064)</td>
<td></td>
</tr>
<tr>
<td>HostCountry*OwnEthnic</td>
<td></td>
<td>0.0306**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0109)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.5752***</td>
<td>0.5682***</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.0109)</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Technology FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.0191</td>
<td>0.0195</td>
</tr>
<tr>
<td>Focal Patents</td>
<td>8939</td>
<td>8939</td>
</tr>
<tr>
<td>Observations</td>
<td>262126</td>
<td>262126</td>
</tr>
</tbody>
</table>

Standard errors clustered at the focal patent level in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001
### Table 8. Knowledge sourcing and protection in the host country - LPM regression

<table>
<thead>
<tr>
<th>Dep. Var. = Citation</th>
<th>Non protected</th>
<th>Protected in non-OECD country</th>
<th>Protected in OECD country</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Foreign</strong></td>
<td>-0.1332***</td>
<td>-0.1526***</td>
<td>-0.1488***</td>
</tr>
<tr>
<td></td>
<td>(0.0034)</td>
<td>(0.00212)</td>
<td>(0.0373)</td>
</tr>
<tr>
<td><strong>HostCountry</strong></td>
<td>0.2421***</td>
<td>0.2309***</td>
<td>0.3667***</td>
</tr>
<tr>
<td></td>
<td>(0.0059)</td>
<td>(0.0076)</td>
<td>(0.0543)</td>
</tr>
<tr>
<td><strong>OwnEthnic</strong></td>
<td>0.0032***</td>
<td>-0.0023</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0047)</td>
<td>(0.0163)</td>
</tr>
<tr>
<td><strong>Foreign*OwnEthnic</strong></td>
<td>0.0498***</td>
<td>0.0143</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>(0.0024)</td>
<td>(0.008)</td>
<td>(0.0053)</td>
</tr>
<tr>
<td><strong>HostCountry*OwnEthnic</strong></td>
<td>0.0369***</td>
<td>-0.0610</td>
<td>-0.0119</td>
</tr>
<tr>
<td></td>
<td>(0.0110)</td>
<td>(0.1234)</td>
<td>(0.0230)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.5254***</td>
<td>0.5763***</td>
<td>0.5763***</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.0102)</td>
<td>(0.0106)</td>
</tr>
<tr>
<td><strong>Year FE</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Firm FE</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Technology FE</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

| R²                   | 0.0192        | 0.0197                       | 0.0192                    |
|                      | 0.0260        | 0.0260                       | 0.0261                    |
| **Focal Patents**    | 8741          | 8741                         | 8741                      |
|                      | 8741          | 8741                         | 8741                      |
| **Observations**     | 252016        | 252016                       | 252016                    |

Standard errors clustered at the focal patent level in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001
Figure 1. Average marginal effects of the main regressors at different values of OwnEthnic (baseline specifications)
Figure 2. Average marginal effects of the main regressors at different values of \textit{OwnEthnic} (including host-country -left- and ethnic base -right- fixed effects)
Figure 3. Average marginal effects of the main regressors at different values of *OwnEthnic*