I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12994

Migrant Inventors and the Technological Advantage of Nations

Dany Bahar
Prithwiraj Choudhury
Hillel Rapoport

FEBRUARY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12994

Migrant Inventors and the Technological Advantage of Nations

Dany Bahar
The Brookings Institution, Harvard CID, CESifo and IZA

Prithwiraj Choudhury
Harvard Business School

Hillel Rapoport
Paris School of Economics, CEPII, CESifo and IZA

FEBRUARY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Migrant Inventors and the Technological Advantage of Nations*

We investigate the relationship between the presence of migrant inventors and the dynamics of innovation in the migrants' receiving countries. We find that countries are 25 to 60 percent more likely to gain advantage in patenting in certain technologies given a twofold increase in the number of foreign inventors from other nations that specialize in those same technologies. For the average country in our sample, this number corresponds to only 25 inventors and a standard deviation of 135. We deal with endogeneity concerns by using historical migration networks to instrument for stocks of migrant inventors. Our results generalize the evidence of previous studies that show how migrant inventors "import" knowledge from their home countries, which translates into higher patenting in the receiving countries. We interpret these results as tangible evidence of migrants facilitating the technology-specific diffusion of knowledge across nations.

JEL Classification: O31, O33, F22

Keywords: innovation, migration, patent, technology, knowledge

Corresponding author:
Dany Bahar
The Brookings Institution
1775 Massachusetts Ave NW
Washington DC, 20001
USA

E-mail: dbahar@brookings.edu

* The authors are thankful to Ina Ganguli, Francesco Lissoni, Ernest Miguelez, and James Sappenfield, as well as participants at the Wharton Conference on Migration, Organizations, and Management (2019) for insightful comments. We also thank four anonymous referees for comments and suggestions. Bahar acknowledges financial support for this research provided by CAF-Development Bank of Latin America to the Brookings Institution. All errors are our own.

"Through the ages, the main channel for the diffusion of innovations has been the migration of people."

(Cipolla, 1976, p. 121)

1 Introduction

It is a known fact that German and Austrian Jewish scientists and inventors who fled from Nazi Germany during the mid 1930s played a crucial role in boosting the innovation capabilities of the countries that received them and, in particular, of the United States. Moreover, this boost in innovation–which was a result of higher combined patenting activity for both immigrants and natives–was in research fields (such as chemistry) where German and Austrian scientists were active inventors in their home countries prior to the war (Moser et al., 2014). While there is plenty and growing evidence of the impact of migration on innovation (e.g., Kerr, 2008; Agrawal et al., 2008; Breschi and Lissoni, 2009; Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010; Agrawal et al., 2011; Freeman and Huang, 2015; Ganguli, 2015; Bosetti et al., 2015; Choudhury, 2016; Akcigit et al., 2017; Breschi et al., 2017; Bernstein et al., 2018; Miguélez, 2018; Choudhury and Kim, 2018; Doran and Yoon, 2019; Burchardi et al., 2019; Miguelez and Noumedem Temgoua, 2019)[1], there is less systematic evidence–based on a larger number of countries examined over a period of time–documenting the role migrant inventors and scientists play in their receiving countries in terms of boosting innovation in those same technologies. This paper attempts to fill this gap in the literature.

In particular, we ask: Do migrants boost patent production in their countries of destination (alt. origin) in the same technology classes in which their home (alt. receiving) countries specialize? We find that, for any given country c and technology p, a twofold increase in the number of migrant inventors from other nations that specialize in patenting in technology p is associated with a 25% to 60% increase in the probability that c gains global technological advantage in p within a decade. In our exercise, gaining technological

[1] See Lissoni (2018) for a comprehensive survey of this literature.

advantage implies achieving a number of patent applications in technology p that is proportionally larger than the global average. This twofold increase, for the average country in our sample corresponds to only about 25 inventors (and a standard deviation of 135). Our research question builds on the findings by Bahar and Rapoport (2018) –who claim that migrants induce industry-specific productivity shifts (as measured by export dynamics)– by investigating the link between migrant inventors and innovation dynamics as one plausible mechanism driving their results.

Our paper attempts to integrate two previously disconnected yet important strands of the innovation and patenting literature: the literature on comparative patenting across countries[2] and the literature documenting the role of migrant inventors in facilitating knowledge production across borders. The importance of geographic and political borders for knowledge transfer and production has been long studied in the patenting and innovation literature. Building on the rich literature of geographic localization of knowledge spillovers (Thompson and Fox-Kean, 2005; Henderson et al., 2005), Singh and Marx (2013) find a strong role of political borders in knowledge diffusion: The authors find both country and state borders to have independent effects on knowledge diffusion beyond what just geographic proximity in the form of metropolitan collocation or shorter within-region distances can explain. In this literature, Foley and Kerr (2013) find that increases in the share of a firm's innovation performed by inventors of a particular ethnicity are associated with increases in the share of that firm's affiliate activity in countries related to that ethnicity. Almeida et al. (2014) study patent data and find that the utility of ethnic knowledge and collaborators depends on

[2]For instance, Glismann and Horn (1988) present an analysis of invention performance, as measured by patenting activities of six countries (France, Italy, Japan, United Kingdom, USSR, West Germany) relative to the United States for 41 SIC industries over 1963-1983, suggesting the existence of "catching-up" processes in terms of patenting activity. More recently, De Noni et al. (2018) assert that less innovative European regions (referred to as 'lagging-behind regions' in their paper) must actively work to reduce the gap between them and knowledge-intensive regions. The authors employ a seven-year panel dataset (2002–2008) using patent data at a regional level to validate the hypothesis that collaborations, and specifically with highly innovative regions, positively affect the innovation performances of lagging-behind regions.

the level of inventor embeddedness in the community. In a recent paper, Kerr and Kerr (2018) connect collaborative patents to the ethnic composition of the U.S. inventors and cross-border mobility of inventors within the firm. In another recent study, Berry (2018) studies global patent production within multinational firms and finds that "knowledge network embeddedness" with the headquarters, host country, and other countries increases future patent production for MNEs. Choudhury and Kim (2018) exploit a natural experiment and supply shock of Chinese and Indian migrant inventors in the U.S. to find that ethnic migrant inventors are instrumental in transferring contextual knowledge (i.e., knowledge locked in geographic regions), such as the knowledge of herbal medicine, across borders.

Our paper finds a robust pattern of migrant inventors impacting cross-country innovation dynamics for particular technologies with rich patenting activity in their home countries prior to their move. In that sense, our findings generalize some of the important findings by Moser et al. (2014) on the spike of innovation in chemistry-related fields due to the inflow of Jewish scientists and inventors to the U.S. in the early 1930s (summarized above); as well as findings by Bernstein et al. (2018), who study patenting behavior of immigrant inventors to the United States in recent decades and find that these inventors tend to "import" foreign technologies into the U.S. (which they measure by the higher propensity of these migrant inventors to cite foreign patents and to work with foreign inventors). Beyond studies that focus on particular countries or historical episodes, our paper–to the best of our knowledge–is the first to use contemporaneous data to establish at a global scale that migrant inventors do shape technology-specific innovation dynamics. This is what we consider the main contribution of our study.

We arrived at these findings by linking and analyzing several sources of data for 95 countries around the globe. First, we use data from the OECD on patenting activity reported by the United States Patenting Office (USPTO) for 651 technology subclasses as defined by the International Patenting Classification (IPC). Our focus on particular technologies is aligned with a rich prior literature in innovation that has used classification of patents according to technologies to study knowledge relatedness and technological distance be-

tween countries (e.g., Jaffe, 1986, 1989; Breschi et al., 2003).[3] To our measure of innovation based on patent classification, we incorporate data on bilateral stocks of migrant inventors compiled by Miguelez and Fink (2017), which measures the presence of foreign inventors in every host country. As will be described in detail later, we use these data to study the relationship between the international mobility of inventors and the spike in patenting activity in their receiving countries, in particular technologies in which their home countries have a technological advantage. To measure this, we employ the Revealed Technological Advantage (RTA) measure, based on Soete (1987). For each country, technology, and year in our sample, we quantify its RTA and use it to measure the yearly intensity with which a country specializes in a given technology. For any technology, an RTA above 1 implies that the inventors in a country in a given year filed proportionally more patents than the world as a whole.

Using two decade-long periods (1990-2000 and 2000-2010), our exercise looks at two different outcomes to measure the dynamics of specialization of a country. First, we construct a binary variable that takes the unit value if a country-technology pair achieved an RTA of 1 or more in a period of 10 years, conditional on that country having started off the decade with zero patent applications in that same technology. We refer to this phenomenon as a technological "take-off." Second, in order to study accelerations, we calculate the decade-long growth rate in the number of patent applications for each country-technology pair (which naturally is defined only for country-technology pairs with some patent activity in the baseline period).[4] We then proceed to explore the extent to which the presence of migrant inventors from (alt. to) countries that have a technological advantage in a specific technological class explains the take-off and acceleration of that same technology

[3]Specifically, Breschi et al. (2003) employ a measure of knowledge-relatedness, using co-classification codes contained in patent documents, and examine the patterns of technological diversification of the whole population of firms from the United States, Italy, France, UK, Germany, and Japan patenting to the European Patent Office from 1982 to 1993.

[4]These two dependent variables are consistent with some of our previous work focused on measuring dynamics of comparative advantage based on international trade data (e.g., Bahar et al., 2014; Bahar and Rapoport, 2018; Bahar et al., 2019).

in their receiving (alt. sending) countries over the course of the following decade.

In order to deal with endogeneity concerns arising from migrant inventors choosing their destination based on private information, the existence of previous trends on technology-specific patent production, or the presence of any other omitted unobservable variable that could bias our estimates, we make use of two sets of instrumental variables (IVs): 30-year-old historic migrant networks as well as the predicted number of migrant inventors based on push and pull factors.[5] We use these two variables, separately, to instrument for the presence of inventor migrants from the same nationalities. In order to rule out the possibility that our results are being driven by prior trends not related to the actual presence of migrant inventors, we perform a number of falsification tests that make our main results disappear. We complement this with a number of additional robustness tests to deal with possible alternative explanations to our results, which we explain in detail below.

The rest of the paper is structured as follows: Section 2 outlines our empirical strategy; Section 3 summarizes our main results; Section 4 conducts subsample analysis and summarizes results from several robustness checks; Section 5 concludes. There is also an Online Appendix that accompanies the paper.

2 Empirical strategy

2.1 Research question

We investigate the relationship between international migration flows and the dynamics of the innovation in migrants' receiving and sending countries. This question follows a research agenda exemplified in Bahar and Rapoport (2018), which explore the role of migration–generally defined–on comparative advantage dynamics using exports data. The main conclusion from that

[5] As will be explained in detail later, our second IV approach is based on Card (2001) and constructed along the same lines as Burchardi et al. (2018) and Burchardi et al. (2019).

study is that migrants serve as drivers of knowledge diffusion, which is reflected in the ability of countries to become significant exporters of the same goods that the migrants' countries of origin specialize in.

The results by Bahar and Rapoport (2018), suggestive of migrant-driven knowledge diffusion, lack specificity in terms of the underlying channels through which this process occurs. In this study, we shift our focus to innovation dynamics and the role that migrant inventors–a particular subset of high-skilled migrants–play in it. We are interested in whether countries' ability to innovate in specific technologies (without prior patenting activity) is influenced by the presence of migrant inventors. Specifically, we ask the following question: Can migrants induce patenting activity in their receiving (alt. sending) countries in the same technologies that their home (alt. destination) countries have an advantage in? For the sake of better understanding, let us use a simple example. Suppose there are two countries in the world: Israel (a country that specializes in patenting water technologies) and Chile (a country that specializes in patenting mining technologies). The analogous question then becomes whether the presence of more Israelis in Chile can explain its specialization in water technologies and whether this same presence is also associated with the ability of Israel to specialize in mining technologies, as measured by patent applications.

2.2 Main data sources and sample construction

Data on patent applications (which we also refer to as patent production throughout the paper) come from the OECD Stat database (OECD, 2014). It counts all patents applications registered by the U.S. Patent and Trademark Office (USPTO) by country of the inventor(s). The count disaggregates the number of patents for each technology subclass based on the International Patent Classification (IPC). An IPC subclass is defined by four characters, letters, and numbers. Throughout the paper, whenever we refer to a technology, we are referring to an IPC subclass (which we often refer to as IPC code). The original dataset covers patenting of 121 countries, and it extends from years 1976 to 2011. The assignment of patents to countries is based

on the declared residence of the inventor(s) of the patent.[6] The dataset also includes figures for patents granted by the USPTO, also per country and IPC code; as well as all patent applications to and granted by other patent offices or treaties, such as the European Patent Office (EPO) and the Patent Cooperation Treaty (PCT), which we also incorporate in our analysis for robustness checks.

Our baseline specification, however, uses patent applications to the USPTO unless otherwise noted, given its more ample coverage of patenting activity.[7] In fact, Figure 1 plots the number of patent applications by year and source for all three–USPTO, EPO, and PCT. The figure shows that the USPTO accounts for a significantly larger number of patent applications than the other sources, with about 300,000 applications in 1990, 600,000 in 2000, and nearly 900,000 in 2010. While EPO is the second-largest source of patent applications for years 1990 and 2000, in our sample, PCT overpasses it in year 2010.

[Figure 1 about here.]

We also limit our main results to patent applications, as opposed to granted patents. This is because a patent is typically granted a few years after the invention actually happened. Hence, patent applications better fit our purposes of measuring production of innovation in a given year. This is consistent with the data, as portrayed in Figure 2 plotting total USPTO patents applications and grants for years 1990, 2000, and 2010. As expected, the number of applications surpasses the number of grants.

[Figure 2 about here.]

Our second main source of data is bilateral international stock of inventors compiled by Miguelez and Fink (2017). The dataset measures for every

[6] For the (relatively few) cases of global collaborative patents (e.g., patents with inventors residing in different countries), the dataset assign a patent to each one of the countries of the inventors.

[7] In Online Appendix Section G we present results using EPO and PCT figures that are robust to our baseline estimations.

pair of countries, c and c', the number of patents by inventor from country c' in country c and vice versa. It is based on patent applications filed under the PCT and has data for about 200 countries. The figures in this dataset are an imperfect measure of the stock of foreign inventors in each country by nationality of the inventor and year. They are imperfect because the number of inventors is contingent on their patenting activity. In other words, it could count the same foreign inventor multiple times or, on at the other extreme, ignore her in a given year. For instance, if a foreign inventor living in country c files two patents in year t, she would be double counted. However, if a foreign inventor living in country c has patenting activity in years $t-1$ and $t+1$ but not in year t, then she would not be accounted for in the data in year t. To overcome this possible fluctuation, we compute the average stock from 1981 to 1990 of inventors living in each country c from each country c' as our measure for 1990, and we compute the average stock from 1991 to 2000 for our measure of inventor migrants in year 2000. While this is not a perfect solution, the average would not be driven heavily by particular outliers in the data. Despite these important caveats, we refer to these numbers throughout the paper as the stock of inventor migrants.

We include in our main dataset other bilateral measures to use as baseline controls: FDI stocks as well as data on bilateral trade. The FDI data comes from the OECD Stat database and tracks FDI flows to or from OECD member countries (thus, it also reports FDI for non-OECD as long as it is to or from an OECD partner).[8] Using these data, we compute FDI stocks for the periods 1985 to 1990 and 1991 to 2000. We also use bilateral trade data that come from UN Comtrade with corrections implemented by Hausmann et al. (2014). With this dataset, we compute stocks of bilateral trade for the periods 1985 to 1990 and 1991 to 2000 to be used as baseline controls. Both the FDI and trade flows are deflated using the U.S. GDP deflator (base year 2000) from the World Development Indicators (WDI) by the World Bank

[8]The FDI data by the OECD includes all financial flows that are cross-border transactions between affiliated parties (direct investors, direct investment enterprises and/or fellow enterprises) recorded during the reference period. The main financial instrument components of FDI are equity and debt instruments.

before being transformed into stocks.[9]

We complement our dataset with overall bilateral migration from Ozden et al. (2011), which we use as part of our identification strategy. The migration dataset consists of total bilateral working age (25 to 65 years old) foreign-born individuals for years 1960, 1970, 1980, 1990, and 2000.

The final sample resulting from merging all the different datasets described above includes figures on patent applications and on migrant inventors for 95 countries across 651 different technology subclasses (i.e., four-character IPC codes). The list of countries with relevant statistics is presented in Online Appendix Section A. The final number of countries is a result of limiting the sample to only countries with some patenting activity in any technology subclass and any presence of migrant inventors. In order to measure decade-long changes in patenting activities for country-technology pairs, we define two decade-long periods (1990-2000 and 2000-2010) for the analysis.

2.3 Empirical strategy

The aim of the paper is to study dynamics in patent production by a country in a particular well-defined technology subclass (measured by patent applications) as a function of the presence of foreign inventors from countries that specialize in that same technology. To do so, we need a measure to quantify the extent to which a country specializes in a particular technology. Our choice is the Revealed Technological Advantage (RTA) index, based on Soete (1987) which, in turn, is analogous to the Revealed Comparative Advantage (RCA) index by Balassa (1965) that is used in international trade.

We compute the RTA for each country and technology subclass in a given year as follows:[10]

[9] We use 1985 as the lower limit for calculating these stocks given source data limitations.

[10] This is analogous to how RTA is computed in the dataset by OECD (2013), though we construct the index ourselves.

$$RTA_{c,p} \equiv \frac{patents_{c,p} / \sum_{p} patents_{c,p}}{\sum_{c} patents_{c,p} / \sum_{c}\sum_{p} patents_{c,p}},$$

where $patent_{c,p}$ is the number of patent applications by inventors in country c in technology subclass p. This is an annual measure. For example, in the year 1990, about 3.25 percent of all patents applications by *Austrian* inventors belonged to technology subclass A63C, which corresponds to "Skates, skis, water-shoes; roller skates; courts; and rinks." Overall, patent applications that year in that same technology by inventors from *all over the world* represented 0.14 percent of all patents applications. Hence, Austria's RTA in technology A63C in the year 1990 was $RTA_{AUT,A63C} = 3.25/0.14 \approx 23$. This means that inventors in Austria patent 23 times more in technology A63C than the world as a whole.

We believe using RTA to measure patenting intensity is proper for several reasons. First, RTA allows us to measure the specialization of one country in a particular technology with respect to the rest of the world, not with respect to another single country. Second, our measures of patent applications are all based on a single patent agency and, thus, the numbers are comparable across countries and years. Third, similar to Balassa's RCA, the benchmark value of 1 and above has an intuitive meaning, as can be understood through the above example.

To study the question at hand with our sample, we follow the empirical specification by Bahar and Rapoport (2018) and estimate:

$$\begin{aligned} Y_{c,p,t\to T} &= \beta_{im} \sum_{c'} inventors^{im}_{c,c',t} \times R_{c',p,t} + \beta_{em} \sum_{c'} inventors^{em}_{c,c't} \times R_{c',p,t} \\ &+ \beta_{FDI} \sum_{c'} FDI_{c,c',t} \times R_{c',p,t} + \beta_{trade} \sum_{c'} trade_{c,c',t} \times R_{c',p,t} \\ &+ \gamma Controls_{c,p,t} + \alpha_{c,t} + \eta_{p,t} + \varepsilon_{c,p,t} \ , \end{aligned} \tag{1}$$

where c represents a country, p represents a technology subclass (i.e., IPC code), and t is a time subscript (T is another time subscript such that $T > t$). The definition of the dependent–or left-hand side (LHS)–variable $Y_{c,p,t\to T}$, changes with the estimation of different outcomes that measure

changes in the intensity of patenting of a country in a given technology. The first outcome we use is a binary variable, which we refer to as a technological "take-off. It measures cases when a country with no patent applications whatsoever in a given technology at time t gains technological advantage in that same technology at time $T = t + 10$. In this case, we define $Y_{c,p,t\to T}$ as a binary variable equal to 1 if the number of patent applications in country c and technology p results in having an RTA of 1 or more in the period of time between t and T, conditional on having zero patent applications in that same technology at the beginning of the period. That is:

$$TakeOff_{c,p,t\to T} = 1 \text{ if } patents_{c,p,t} = 0 \text{ and } RTA_{c,p,T} \geq 1.$$

We impose two additional conditions on our take-off measure to prevent our results from being driven by noise. First, the country-technology pair under consideration must keep its RTA value above 1 for four years after the end of the year T (e.g., have a minimum RTA of 1 during the years $[T, T + 5]$). Second, the country-technology pair under consideration must have had an RTA value equal to 0 during all four years before the beginning of year t (e.g., have a maximum RTA of 0 during the years $[t - 5, t]$).

We alternate our LHS variable with a measure of growth in patent applications for every country and technology subclass from years t to T. In that case, $Y_{c,p,t\to T}$ is simply the annual compound average growth rate (CAGR) in the number of patents in technology p granted to inventors in country c from years t to $T = t + 10$, conditional on having more than zero patent applications in that same technology at the beginning of the period. That is:

$$CAGR_{c,p,t\to T} = \left(\frac{patents_{c,p,T}}{patents_{c,p,t}}\right)^{1/T-t} - 1 \text{ if } patents_{c,p,t} > 0.$$

Our main variables of interest are denoted by $\sum_{c'} inventors^{im}_{c,c',t} \times R_{c',p,t}$, and $\sum_{c'} inventors^{em}_{c,c',t} \times R_{c',p,t}$, where $R_{c',p,t} = 1[RTA \geq 1]$. These variables can be interpreted, respectively, as the stock of immigrant inventors from and of emigrant inventors to other countries (denoted by c') at time t, that

specialize in the production of patents classified under technology subclass p, as indicated by the dummy variable $R_{c',p,t}$.

As controls, we also include the sum of the stock of FDI (inflows plus outflows) and the sum of the stock of trade (imports plus exports), using the same weighting structure as above. Including these controls allows us to reduce omitted variable bias when estimating β_{im} and β_{em}. This is because trade and/or capital flows with the same countries where the inventors come from or go to (denoted as c') could also explain innovation dynamics in country c. Using the same weighting scheme allow us to control for the total trade and FDI to those same countries c'. While ideally we would include trade and investment from those countries that relate to each particular technology p, such data is not available. Since our weighting procedure accounts for total trade and FDI from those countries, they are inclusive of flows that relate to particular technologies. Thus, we believe that measures control for these plausible channels.[11]

In addition, we include country-year fixed effects, denoted as $\alpha_{c,t}$, to control for any country-level time-variant characteristics that correlate with both national migration determinants and aggregate productivity levels, such as income, size, institutions, etc. $\eta_{p,c}$ represents technology-year fixed effects, to allow for a different constant for each combination of year and IPC technology subclass.

We also include a vector of controls for baseline variables when measuring using CAGR on the LHS: the baseline (initial) level of patent applications for that same technology in that country, as well as the previous period CAGR of patent applications in the same technology to control for previous trends. To avoid undetermined lagged growth rates when the initial level of patents is zero we add 1 to both the number of patents in the numerator and the

[11] Of course, there is a genuine discussion to have on whether these are, in fact, "bad controls". In other words, if the inflow of inventors from countries that specialize in a certain technology subclasses triggers more trade and investment that in turn boost innovation dynamics in the receiving country, then we would be underestimating our overall effect. However, we decide to keep them in the baseline specification as we are interested in estimating β_{im} and β_{em} which would measure the partial correlation (or marginal effect) regardless of trade and investment. However, in Online Appendix Section E we show that our results are robust to excluding these controls.

denominator. Thus, we also add as a control a binary variable indicating whether $patents_{c,p,t-10} = 0$ (at the beginning of the previous period, i.e., 1980 or 1990).

All level variables are transformed using the inverse hyperbolic sine (MacKinnon and Magee, 1990). This linear monotonic transformation behaves similarly to a log-transformation, except for the fact that it is defined at zero. The interpretation of regression estimators in the form of the inverse hyperbolic sine is similar to the interpretation of a log-transformed variable.[12] Thus, since our LHS variables are not transformed using a logarithmic scale, the interpretation of the estimators are linear-log.

2.4 Identification

Our main goal is to get unbiased estimators for β_{im} and β_{em}. This is challenging, as one might expect that the choice of country for foreign inventors might be correlated with dynamics of specialization in certain technologies. In other words, there might be other country-technology-time characteristics, perhaps unobservables, that can explain both the inflow and outflow of inventors and dynamics of patent production. We try to overcome this by estimating our specification using 2SLS, using two sets of instrumental variables (IVs) for both immigrant and emigrant inventors (e.g., a set of IVs, in our case, include two variables–one for each endogenous regressor).

First, for each country-technology-year combination, we instrument the immigrant and emigrant inventors with the total stock, lagged by 30 years, of immigrants from and emigrants to those same countries (e.g., we apply the same weights defined by$R_{c',p,t}$). In that sense, our instruments are $\sum_{c'} immigrants_{c,c',t-30} \times R_{c',p,t}$ and $\sum_{c'} emigrants_{c,c',t-30} \times R_{c',p,t}$.

Second, we construct an additional set of IVs for the number of immigrant and emigrant inventors using pull and push factors, computed with the data itself, expanding the approach first introduced by Card (2001). Recently, Burchardi et al. (2018) and Burchardi et al. (2019) use a similar approach to

[12] The inverse hyperbolic sine ($asinh$) is defined as $log(y_i + \sqrt{(y_i^2 + 1)})$. Except for small values of y, $asinh(y_i) = log(2) + log(y_i)$.

estimate the effect of immigration on FDI and innovation, respectively, for the U.S. Our instrument is based on a prediction of the actual stock of inventor migrants between c and c', combining a "push" and "pull" components as follows:

$$invIV_{c,c',t} = push^{-c}_{c',t} \times pull^{-c'}_{c,t} \times \sum_{c} \sum_{c'} inventors_{c,c',t} \ ,$$

where $push_{c',t}$ is the share of all migrant inventors in year t *from* country c' to all other countries, and $pull_{c,t}$ is the share of all migrant inventors in year t *to* country c from all other countries.

The superscripts $-c$ in the push factor and $-c'$ in the pull factor are there because –in order to further reduce any endogeneity concerns– we exclude from the calculation the bilateral flow of migrant inventors from and to the country of the corresponding observation. For clarity, the terms $push^{-[c]}_{c',t}$ and $pull^{-[c']}_{c,t}$ are computed as follows:

$$push^{-c}_{c',t} = \frac{\sum_i inventors_{i,c',t}}{\sum_i \sum_{c'} inventors_{i,c',t}}, \text{ where } i \neq c$$

$$pull^{-c'}_{c,t} = \frac{\sum_j inventors_{c,j,t}}{\sum_c \sum_j inventors_{c,j,t}}, \text{ where } j \neq c'$$

Where c and c' are a receiving and a destination country respectively, while i and j are their respective partner countries for each bilateral flow.

Finally, to construct our instruments, we apply the same weighting scheme explained above using the *predicted* number of inventors between each pair of countries c and c' based on pull and push factors, as: $\sum_{c'} invIV^{im}_{c,c',t} \times R_{c',p,t}$ and $\sum_{c'} invIV^{em}_{c,c',t} \times R_{c',p,t}$.

Both set of instruments, in order to be valid, should be able to explain enough variation in the endogenous variables. We expect this to be the case because historic migrant communities should work as a pull factor for the decision of inventors to migrate to particular countries. In addition, the predicted number of inventors based on push-pull factors should be a good enough predictor of actual migrant inventor flows between countries. We find

this to be the case in our sample based on the reported first-stage statistics[13] (with some exceptions, which are discussed thoroughly).[14]

In addition to the explanatory power of the first stage, for our instruments to be valid (and, thus, to be able to interpret our 2SLS estimators as causal), they need to comply with the *exclusion restriction*. In our case, this exclusion restriction can be verbalized as follows: It must be *technology-specific* production (e.g., patent applications) in any given country and not correlated with our instruments, other than through the presence of inventor migrants today. Furthermore, to be able to interpret our 2SLS estimators as causal, we also must assume that countries do not engage in *technology-specific* innovation agreements based on their historic migrant networks that are not captured via FDI or trade (since we are controlling for those flows, too).[15]

In the case of our first instrument, the assumption is that the existence of a historic migrant community from country c' in the destination country explains the flow of migrant inventors, a particular subset of high-skilled migrants; but–at the same time–that historic migrant community from country c' does not explain future dynamics of patent production other than through the inventor migrants who had arrived later. While we believe this is a reasonable assumption, one might think that the presence of a historic migrant community might affect the diffusion of knowledge through channels unrelated to the presence of migrant inventors (and, also, unrelated to trade and

[13] In all of our 2SLS estimations we report the Kleibergen-Paap F statistic to be used to determine whether instruments are weak, which according to Stock and Yogo (2005), must be above 16.78 when using two endogenous variables and two instruments. We acknowledge that these critical values are not strictly usable in the case when we do not assume i.i.d., but for the most part, unless otherwise noted, our Kleibergen-Paap F statistics are high enough that there are no reasons for concern regarding weak instrumentation.

[14] Naturally, there could be concerns that the observed first-stage correlations between migrant inventors and the instruments are artificially being driven by the weighting scheme. However, this is not the case: both instruments have a strong explanatory power on current stock of migrant inventors before we apply the weights and transform them into country-technology-year variables. We include evidence of this in Online Appendix Section B.

[15] Note that since we include trade and FDI in our 2SLS estimation (i.e., they are also part of the first stage) we already control for the fact that the instruments might affect future innovation through FDI and trade.

capital flows with their countries of origin). If this is the case, then our instrument would be invalid.

Therefore, we also present results with another instrument–the predicted number of inventor migrants using push and pull factors–in order to further validate our estimations. In the case of our second instrument and in order to interpret our estimates as causal, our assumption is that push and pull factors used to predict inventor migrant flows are not correlated to innovation dynamics other than through the presence of the migrant inventors themselves. Since we constructed the push and pull factors in a way that excludes information about the country-pair under consideration, we believe this assumption is reasonable. Note, too, that in this setting, the inclusion of country-year fixed effects controls for the overall attractiveness of inventors to the country under consideration, further reducing endogeneity concerns. Thus, in a sense, we are exploiting mostly the push factor, which is more likely to be exogenous to innovation dynamics for a country-technology pair.

In summary, we believe these are reasonable assumptions to make, though we acknowledge there might still be weaknesses in our approach. Thus, to complement our efforts in establishing the relationship, we also perform a number of falsification tests showing that our results are indeed driven by the flow of inventors and do not respond to previous trends or other variables (observables or not) that are not accounted for in our main estimation.

2.5 Descriptive statistics

Table 1 presents descriptive statistics for our sample. Panel A presents the summary statistics for the subsample that focuses on technology take-offs (i.e., for all observations of c, p, and t for which $RTA = 0$), while Panel B shows the same for the subsample focusing on growth of patent production (i.e., for all observations of c, p, and t for which $patents_{c,p,t} > 0$).

[Table 1 about here.]

Panel A of Table 1 describes that the unconditional probability of a take-off for the average country and average technology subclass, pooling

observations for two decades (1990 to 2000 and 2000 to 2010), is 2.2 percent. Note that this is based on the sample limited to country-technology pairs with zero patent applications at the initial year of each decade. Panel B presents statistics based on the complementary sample; that is, with at least one patent application at the beginning of each decade for every country-technology pair. This sample is used to measure the impact of migration on growth of technologies in terms of patent applications. The 10-year CAGR for the average country-technology pair–also pooling observations for two decades–is 0.7% and varies from -30% to 80% in both extremes for some country-technology pairs. In this sample the baseline number of patent applications for the average country-technology pair is about 16.45. Notably, the number of observations that make up the "technology take-off" sample is almost six times as large as the sample described in Panel B. This is not surprising, since the vast majority of country-technology pairs have, in fact, no patent activity.

The tables also include figures for immigrant and emigrant inventors weighted using the scheme used on the right-hand side of Specification (1). According to Panel A, which focuses on take-offs, the average country-technology pair in the sample has about 24 inventors who have immigrated from countries that specialize in that same technology and about 80 inventors who have emigrated to countries specializing in that technology. Those same figures in the sample summarized in Panel B are about 710 and 610, respectively. The larger numbers of average inventor migrants in Panel B responds to the fact that such a sample is composed mostly by developed nations, which host many more inventors (as those include only country-technology pairs with some patent applications).

The table also summarize our IVs. The first set of instruments includes the 30-year lagged stock of immigrants and of emigrants, weighted using the same weighting scheme as our right-hand side variables of interest. The average values for these figures are about 69,000 immigrants and 142,000 emigrants in Panel A and 499,000 immigrants and 449,000 emigrants in Panel B. As expected, these numbers are significantly larger than the number of inventor migrants, as inventors are only a very small subset of all migrants.

The second set of instruments include the predicted number of immigrant inventors and emigrant inventors based on the pull and push factors, using the same weighting scheme as the actual stock of immigrant and emigrant inventors. In Panel A, the average figures for these IVs are 18 predicted immigrant inventors and 55 predicted emigrant inventors, whereas in Panel B, the corresponding statistics are 469 predicted immigrant inventors are 403 predicted emigrant inventors. The same reasoning as before with respect to the larger average numbers seen in Panel B as compared to Panel A applies in this case.

Finally, the table also has subsample statistics on total trade and FDI, in billions of dollars, constructed using the same weighting scheme.

3 Main Results

The main question we aim to answer is whether a country can become a significant innovator of a particular technology–what we call a technology take-off–if it has immigrant inventors from (or emigrant inventors in) other countries that specialize in patenting activity in that same technology. A simple look at the raw data, represented in Figure 3, presents preliminary evidence of that being the case. Average take-off rates of country-technology pairs are higher whenever they host a larger number of inventor immigrants from other countries that specialize in those same technologies, for both periods 1990-2000 and 2000-2010. In particular, the figure shows the unconditional probability of a country-technology pair taking off in the period 1990-2000 is about 0.2% when it had a stock of immigrant inventors below the median in the baseline year (1990), compared to 0.8%–about four times as much–when the stock of immigrant inventors is above the median. For the 2000-2010 period, the corresponding figures are 2.3% and 5.7%. The figure also shows the same figures for decade-long growth rates for patent applications. During the 1990-2000 period, country-technology pairs with a stock of immigrant inventors (from other countries that specialize in that technology) below the sample median grew at a pace of 0.24% a year, compared to 1.08% a year for those with a stock of immigrant inventors above

the sample median. However, for the 2000-2010 decade, we see the opposite pattern: country-technology with a stock of immigrants below the median grew faster in terms of patent applications than country-technology pairs with a stock of immigrants above the median.

[Figure 3 about here.]

Many confounding factors could explain Figure 3, of course. Therefore, next, we present results using more rigorous estimation techniques.

3.1 OLS and 2SLS estimations

The estimation of Specification (1) is presented in Table 2. The upper panel estimates the changes in the probability of technology subclass take-off as a function of migrant inventors.[16] In the estimations, as mentioned above, all of the regressors have been transformed using the inverse hyperbolic sine and, therefore, the interpretation of the coefficients correspond to semi-elasticities (i.e., linear-log).[17] The first three columns show results using OLS as the estimation technique, whereas Columns 4 to 9 use 2SLS estimations, using two set of IVs, as detailed in Section 2.3.

[Table 2 about here.]

The results of Panel A estimate the partial correlation of our variables of interest: immigrant inventors from and emigrant inventors in countries specializing in a given technology subclass at the beginning of the decade (separately in Columns 1 and 2 and jointly in Column 3) with respect to the take-off of the same technology by a country.[18] Results in Columns 1 and 3

[16] As explained above, we define a technology take-off as cases where a country achieves a RTA of one or more within a decade, starting off from no patenting activity (see Section 2.3 for formal definition).

[17] We refrain from rescaling our right-hand side variables in terms relative to population. This is mainly because the inclusion of country-year fixed effects control for population size, reducing concerns that our results are driven by scale effects. Yet, the falsification tests we present below also rule out the possibility of scale effects driving our results.

[18] Our estimates are robust to using a maximum likelihood estimator given the binary distribution of our dependent variable in Panel A (specifically, the complementary log-log

show that a twofold larger stock of immigrant inventors from countries that specialize in technology p is associated with an increase in the probability of the receiving country specializing in patent applications in technology p of 0.51 percentage points. Given that the unconditional probability of a take-off is 2.2 percent, this represents an increase in the probability of about 23%. The estimator for emigrant inventors is not statistically different from zero when jointly estimated with immigrant inventors (Column 3).

The economic significance of this number is quite large. Given that the stock of immigrant inventors in the sample is about 24 people for the average country-technology pair, and a standard deviation of 135 inventors, a twofold increase implies a relatively small number of inventors. Thus, according to our results, a small number of migrant inventors has significant and large explanatory power on the likelihood the receiving country will gain advantage in a new technology subclass (in which it had no patent activity beforehand).

Columns 4 to 6 in the upper panel of Table 2 replicate the results using a 2SLS estimator; we use instruments based on the 30-year lagged stock of immigrants from (and of emigrants in) the same countries as the migrant inventors (IV1). Given that we have two instruments, we can use them in estimations that include both variables of interest separately (Columns 4 and 5) as well as jointly (Column 6). Columns 7 to 9 present an alternative 2SLS estimation using our second set of instruments: the predicted number of inventors using contemporaneous push and pull components (IV2), as explained in Section 2.4. Similarly, because this set includes two instruments, we can estimate both variables of interest separately or jointly. Note that the table reports the Kleibergen-Paap F statistics for all 2SLS estimations, which are large enough to eliminate any concerns of weak instrumentation.

The 2SLS results are qualitatively similar to the OLS results, but higher

method which is more appropriate for our setting, following Singer and Willett, 2009) . It is also robust to using the methodology suggested by Horrace and Oaxaca (2006), to deal with the possibility of our results being driven by outliers. For more details, see discussion in Online Appendix Section C. In addition, Online Appendix Section D presents results using alternative left-hand side variables, including a more widespread –and less restrictive– measure of take-off: a binary variable that takes the value of 1 whenever a country-technology pair goes from zero patent applications to any number higher than zero.

in magnitude by a factor of 2 to 3. Interestingly, the two different 2SLS estimations, using very different instruments, yield strikingly similar point estimates, reinforcing the validity of our identification strategy. It is somewhat counterintuitive at first, when comparing the OLS and 2SLS estimations, that the point estimate of β^{IM} becomes larger after the instrumentation. If anything, we would expect a positive bias in the OLS estimates, not a negative one, as unobserved forces that lead to more innovation might also pull immigrant inventors (the opposite would happen for emigration, for which the OLS coefficient is statistically indistinguishable from zero). But it could well be that the 2SLS results correct for biases due to measurement error of our endogenous variables (in fact, as explained in Section 2.2, these measures likely suffer from measurement error).[19] By in large, however, our 2SLS results are inconclusive when it comes to understanding whether the OLS bias is positive or negative: Even though the magnitudes of the 2SLS estimates are larger than that of the OLS, the standard errors have also increased and, thus, we cannot reject the hypothesis that both estimates for β^{IM}(as well as β^{EM}) are statistically different, as shown in Figure 4.

[Figure 4 about here.]

Relying on the 2SLS estimates, the effect of a twofold increase in the number of immigrant inventors results in an increase of 50% to 60% in the likelihood of the receiving country specializing in patent applications in the same technologies in which the migrants' home countries specialize.

Panel B of Table 2 estimates the partial correlation of our variables of interest on the growth rate of technology-specific patent applications for the average country, as a function of inventor immigrants from (and inventor emigrants in) other countries that specialize in that same technology. The sample we use is limited to those country-technology pairs for which the

[19] Since our regressors of interest are a compound variable (i.e., the aggregated number of migrant inventors weighted by $R_{c',p,t}$), the larger magnitude of the 2SLS estimates with respect to OLS suggests our results are not driven by the second term (e.g., a convergence effect due to the "technological gap" between the countries), but rather by the first one which we effectively instrument for while keeping unchanged the weights. We thank an anonymous reviewer for making this point.

initial value of patent applications is above zero, as it is not possible to compute growth rates otherwise. But there is another, more fundamental, reason. In essence, this distinction allows us to focus on innovation dynamics for technologies already being patented in the country. In those cases, arguably, there is already a critical mass of inventors with knowledge on that specific technology subclass, as opposed to cases in which there is no prior patenting activity whatsoever, such as country-technology pairs in the subsample used in Panel A. The distinction between an extensive margin (i.e., take-offs) and an intensive margin (i.e., growth) is often used in the international trade literature when studying the composition dynamics of countries' exports baskets.[20]

The results from Panel B present a conclusion qualitatively similar to that in Panel A. The OLS estimations (Columns 1-3) imply that, for the average country, a twofold increase in the number of immigrant inventors from other countries specializing in technology p explains a higher growth rate in patenting activity of the same technology p of 0.34 to 0.45 percentage points per year over the following decade. Also, according to the OLS results, we find that a twofold increase in the number of inventor emigrants in other countries is correlated with annual growth rates in patent activity that are higher by 0.29 to 0.42 percentage points for the technologies of specialization of the receiving countries. Since the decade-long unconditional growth rate for patent applications is 0.7% (see Panel B, Table 1), then the marginal effect for immigrant inventors according to Columns 1-3 corresponds to an increase of up to 65%. Note that in this sample, however, the average number of immigrant inventors is much larger and corresponds to about 700.

Columns 3-9 estimate the same specification using 2SLS, with the two sets of instruments discussed above. However, across all columns, the first-stage statistics are not nearly as large as the ones in Panel A, implying that

[20]In Online Appendix Section D2 we use a measure of growth in patent applications that is defined for country-technology pairs with initial value of zero (i.e., a symmetric percentage change). With this growth measure as the dependent variable we are able to estimate our empirical specification on *all* the sample and find robust results; though the results are driven by the observations with no initial patenting activity. Our results using this aggregated measure, however, are statistically weaker.

there could be a weak instrumentation that makes all the 2SLS estimations in Panel B results invalid and, therefore, we cannot solve endogeneity concerns for this sample.[21] In fact, as opposed to Panel A, the 2SLS point estimates are quite inconsistent across the different estimations. Therefore, while our findings on technology take-offs seem quite robust, we cannot make any claims about growth rates.

The table also reports estimators for our control regressors, namely FDI and trade. The idea behind including these controls is that they could very well correlate with the flow of migrant inventors and, in turn, could explain future innovation dynamics. The point estimates are volatile around the zero value (seldom statistically significant), across the different specifications in both panels. Their inclusion in our model is part of our identification strategy to reduce concerns of biases in our regressors of interest, β_{im} and β_{em}. The lack of a clear relationship between these controls and the dependent variable is intriguing. However, there is a mechanical and straightforward reason for this (lack of) result: the high multicollinearity of these terms with the country-year fixed effects. Despite the fact that they vary across technology subclasses within each country and year, in practice such variation is very small. Note that the regressors are the sum of all trade and FDI to and from countries where migrant inventors are from or in for that particular year. Since most trade and FDI happens between developed nations, which is also where most of the innovation happens across a wide variety of technology subclasses, these terms are very similar to–and, in fact, their correlation is above 0.8 with– total unweighted trade and FDI figures for every country and year (which are, by definition, perfectly multi-collinear with country-year fixed effects). Therefore, the inclusion of country-year fixed effects eliminates most of the variation needed for properly estimating a partial correlation between these controls and the dependent variables. This is consistent with the fact that we find strikingly similar results if we exclude these controls (see Online Appendix Section E). But, even if there was no multicollinearity problem, the estimation of these controls could result in very small estimates.

[21] Note that the difference in the explanatory power of the instrument could be expected, as the samples used in both panels are very different

This is because, plausibly, only a small share of both trade and capital flows (both of which aggregate inflows and outflows) is relevant for innovation and, in particular, for innovation of specific technologies. Ideally, we would have preferred to use more specific controls, such as trade and capital flows *relevant* for the particular technology under consideration. However, not only is that data not readily available, but it is not clear how to construct such measures, if it is possible at all. In order to be conservative, however, we choose to include them in our baseline specification, but refrain from concluding anything about these flows in our empirical setting.

All in all, our main results support the idea that migrant inventors facilitate the spread of ideas reflected in significant patenting activity in technologies in which their home countries' specialize. While in the main body of the paper we use patent applications as our main data source, Table G3 in the Online Appendix (Section G) shows that our main results are robust to using granted patents as opposed to patent applications. Naturally there could be important gaps between the time of the patent application and the time of its acceptance by the USPTO, and that time gap could be problematic in our setting. However, the fact that the results are robust to using granted patents is complementary to our main estimation, as it provides suggestive evidence that the process through which immigrant inventors affect the dynamics of patent applications for a given technology is also reflected in the ability of the receiving country to convert some of those applications into granted patents.

3.2 Falsification tests

While we believe our estimation methods outlined in the previous section deal, for the most part, with plausible endogeneity concerns between the flow of inventor migrants and countries' patenting activity, we acknowledge that there could be violations to the exclusion restriction of our instrumental variable approach. The nature of our macro-level dataset, while is useful to generalize our findings, also poses important challenges to our ability to perfectly identify the relationship we are studying.

In order to deal further with some remaining endogeneity concerns, we propose two tests to explore whether our results are indeed consistent with the possibility of inventor migrants impacting patenting activity of their receiving countries or, alternatively, they are driven by trends in the data, unrelated to migration, for which we are not accounting. For this purpose, we perform two *falsification tests* by altering the right-hand side variables of interest. While not a perfect approach for identification purposes, we consider it useful to show that our results do respond to actual variation in migrant inventors and not to existing pre-trends or other factors omitted from the analysis.

First, we replicate Specification (1), but this time using the weighting parameter $R_{c',p,t} = 1$ if $RTA_{c',p,t} = 0$. That is, we exploit variation in inventors migrating from and to countries c' that had zero patent applications in technology p at time t. Table 3 reports the results.

[Table 3 about here.]

When alternating the right-hand side variables of interest in this way, we find that the results are very different than the ones presented in Table 2. According to Panel A, the presence of migrant inventors from or in countries with no patent applications in technology p does not consistently explain an increase in technology take-offs as compared to our baseline results (Columns 1-3 present OLS estimators and Columns 4-9 present 2SLS estimators using the two sets of instruments). If anything, the negative point estimates suggest that immigrant inventors from and emigrant inventors to countries that do not patent at all in a specific technology *reduce* the likelihood of a country patenting in that same technology during the following decade. The results in Panel B, which focus on growth, are consistent with those in Panel A (note, though, that for the 2SLS estimates, we have a weak first stage) and show that country-technology pairs grow slower in terms of patent applications whenever they host immigrant inventors from other countries that do not patent in those same technologies.

Second, we alternate our right-hand side variables of interests based on number of inventors that are randomly generated and, therefore, do not

reflect the actual number of inventors. In that context, we present results for two models. The first approach, which we title "Random Model 1" randomizes the number of inventors between countries without altering the distribution. In other words, we "reshuffle" the number of migrant inventors between countries randomly, such that the total global figure for (random) migrant inventors in a given year is the same as the actual number. Thus, the distribution and the average for these two variables are exactly the same. The second approach, referred to as "Random Model 2" does not impose any restrictions whatsoever and simply creates a fake number of migrant inventors based on a random number (from 0 to 1, uniformly distributed) to every country pair and year.

Our exercise is based on 500 iterations, which involve reconstructing the dataset each time (e.g., recomputing the actual and random inventor variables by assigning the weighting structure and then collapsing the dataset to create aggregated figures for every country, technology, and year cell). Figure 5 presents density plots of correlation coefficients between the actual number of inventors and each one of the 500 iterations, for both models. Note that these correlations are computed using just country-pair data, before applying the proper transformations to construct the regressors detailed in our empirical specification.

[Figure 5 about here.]

Note that when the randomization is done in a way that preserves the same distributional characteristics of the original variable (left panel in Figure 5), the correlation between the random and actual variable always takes positive values across the 500 iterations, and its distribution is characterized by having a fat-tail. However, when not imposing any restrictions on the generation of a random number (right panel), the correlation coefficients between the random and actual figures for the 500 iterations distribute quasi-normally and around zero, as is expected when correlating with a truly randomly generated variable. Despite this fundamental difference, we present results using these two approaches and find consistent results.

Figure 6 summarizes our results (one marker for each one of the 500 iterations) when using OLS to estimate our main specification but substituting the main variables of interest $\sum_{c'} inventors^{im}_{c,c',t} \times R_{c',p,t}$ and $\sum_{c'} inventors^{em}_{c,c',t} \times R_{c',p,t}$ with the ones that we constructed through randomization for inventor immigrants and emigrants. Note that the reported estimators for β^{IM} and β^{EM} are a result of including both regressors simultaneously in the regression (analogously to Columns 3 of Table 2). The figure also reports the estimator from our baseline specification, using the actual number of migrant inventors, as reported in Table 2 and denoted by a diamond-shaped marker. Whiskers in the figure represent 95% confidence intervals.

[Figure 6 about here.]

Clearly, the results using a random number of inventors across the 500 iterations is extremely noisy in both models and both estimators (immigrants and emigrants). In fact, when focusing on estimates of β^{IM}, in nearly 45% of the iterations, the result is not statistically different from zero for Random Model 1 (despite the fact that the correlation between the actual and random number of migrants is always positive in this model). The corresponding figure for Random Model 2 is just above 60%. When it comes to β^{EM}, for which the OLS estimator using the real number of inventors is statistically insignificant itself, the estimators based on a random number are statistically insignificant for about 85% of the 500 iterations in the first model and close to 70% in the second model.

These two falsification tests are important, as they should alleviate any remaining concerns that our results are being driven by spurious correlations or previous trends. The results are also useful to alleviate concerns that our results are being driven purely by scale effects. This is particularly relevant when focusing on Random Model 1, which keeps unchanged the aggregated scale global flows of migrant inventors.

4 Supplementary analysis

4.1 Heterogeneity of results

In order to study the relationships documented above in more detail, we reestimate Specification (1) across different subgroups of our sample. We do this to understand whether there are different trends across several dimensions and also to explore whether a particular set of observations in the sample is driving the observed overall results. Table 4 summarizes this exercise.

[Table 4 about here.]

The left panel of Table 4 reports OLS estimates both for β_{im} and for β_{em}, while the two other panels report the 2SLS estimates for the same regressors using both set of instruments (as in the previous section). The table presents results of the specification that uses technology take-offs as the dependent variable (thus, country-technology pair observations are limited to having an initial number of granted patents equal to zero). The estimates reported for β_{im} and for β_{em} are based on a specification that includes both of their corresponding regressors simultaneously. The first row uses all observations (the same sample as presented in the upper panel of Table 2).

The rest of the rows present results for different cuts of the sample. Across the board, based on the 2SLS estimates, we find that our results typically hold only for immigrant inventors, not for emigrants, consistent with our findings so far.

Additionally, our results are being driven by both OECD and non-OECD countries alike.[22] Also, the results are particularly driven by the period 2000-2010, the decade for which most of the patenting activity in our sample occurs.

Finally, we divide our sample into eight IPC sections (which correspond to the first "character" of the four-character IPC subclasses used throughout

[22] We count countries in our sample as OECD members of only if they had been such prior to the first period studied (e.g., the classification does not count countries as OECD members if they became so during the 1990s or the 2000s). For a complete list of countries in the sample, including which ones we categorize as OECD members, see Online Appendix Section A.

the paper). These are human necessities (A); performing operations and transporting (B); chemistry and metallurgy (C); textiles and paper (D); fixed constructions (E); mechanical engineering, lighting, heating, weapons, and blasting (F); physics (G); and electricity (H). While our OLS results do show some heterogeneity in the statistical significance of the results for β_{im}, when it comes to the 2SLS results, we do not find any particular technology section driving the results. In particular, using the first set of instruments, we do not find evidence that our aggregate results are driven by any particular technology. This is almost consistent with the results using the second set of instruments where, except for two technologies (fixed constructions and performing operations), we again find that our results are not driven by any technology class in particular.

4.2 Further robustness tests

We perform a number of supplementary analyses for robustness purposes, which we briefly discuss in this subsection. We provide more details in the Online Appendix.

Our right-hand side variables measuring migrant inventors, trade, and FDI are highly multi-collinear, which could raise concerns about our model being misspecified. However, all of our estimations include a large number of fixed effects, so our estimations correct for different scales. Additionally, we have computed the variance inflation factor (VIF) for our main OLS estimation (Column 3 of Table 2, which include technology-by-year and country-by-year fixed effects) to assess the availability of enough independent variation among correlated variables. The mean VIF value is 1.31, which is within the acceptable range. Therefore, multicollinearity does not appear to be an issue.

Also, we reestimate the specification estimating the impact of migrant inventors on technology take-offs based on patent applications using a non-linear estimation, as it is often done for binary outcomes. In particular, we implement the complementary log-log estimator, which is a better estimator than logit or probit if the probability of take-off is small (e.g., there are

many zeros), as is our case (see Singer and Willett, 2009). We also apply the methodology by Horrace and Oaxaca (2006) that corrects for predicted values of the dependent variable outside the 0 to 1 range. For more details, see Online Appendix Section C.

We also find that our results are robust to variations of our left-hand side variables. In particular, it is robust when using a binary variable that does not depend on RTA, as well as a growth rate that is defined when the initial value is zero (allowing us to estimate using the sample without splitting it into two subsamples). See Online Appendix Section D for more details and the results.

Additionally, in Online Appendix Section D, we document a number of other consistent findings. Our results are robust to using patent application data based on the European Patent Office (as opposed to the USPTO) and find results that are qualitatively and quantitatively similar (see Table G1). This robustness test is particularly important because it shows that our results are likely not driven by the "home advantage" effect, given that such bias would naturally be different for USPTO or EPO patents (Criscuolo, 2005).[23] Our results are also robust to using patent application figures based on the PCT, which we also document (see Table G2); this is also the main source out of which the migrant inventor numbers come. We also replicate our results using data on granted patents (as opposed to patent applications) according to the USPTO (see Table G3). While granted patents typically involve an important time gap, we still believe it is relevant to show that our results are robust to using this measure, as a granted patent is indeed a confirmation that the innovation is novel enough. In the context of our results, this is crucial; it suggests our results are not only driven by an uptake on filling patents, but on actual innovation.

We also explore the possibility that our results are driven by intellectual property theft practices (along the lines of the evidence on industrial espi-

[23] An alternative approach could have been to limit our sample to OECD triadic patents, but this raises a number of other difficulties in terms of possible biases, given that–by definition–these patents represent a subsample of all patents, and it is unclear –at first sight– what are the biases driving the selection of patents into that group.

onage by Glitz and Meyersson, 2017). In other words, we look at whether inventor migrants facilitate the spread of technologies through stealing intellectual property (IP) rather than through knowledge diffusion. To some extent, our results–particularly the ones using granted patents in Online Appendix Section G–address this possibility. This is because our data is based on innovations reported by a formal authority (e.g., the USPTO for our baseline results) that, in essence, should deal with IP thefts. In addition, the idea of IP theft should be less of a concern in our specification given that our country-year fixed effects should control for IP protection intensity in each country at an aggregate level (though they do not control for variations of IP protection intensity within a country for different technologies). As an additional robustness test, we reestimate our main specification excluding China from the sample. We do this because China is a country that (1) is large in size and, therefore, represents an important share of migrants and (2) is known to have weaker intellectual property protection. Our results are robust to the exclusion of China and are documented in Online Appendix Section F1.

5 Conclusions

In this paper, we study and provide robust econometric evidence of the role of immigrant inventors in shaping innovation dynamics in their receiving countries. In particular, our analysis shows that–controlling for other means of exchange such as trade and FDI–countries receiving immigrant inventors from other nations that specialize in patenting in technology p are more likely to have important increases in patent applications in that same technology.

Our estimates imply that a twofold increase in the number of inventor immigrants can explain an increase of 25 to 60 percent in the likelihood of gaining technological advantage in the same technology in which the inventors' home countries specialize. In our sample, this number can be as low as 25 inventors for the average country, with a standard deviation of about 135. Our econometric analysis includes the use of IVs as well as a number of falsification tests to rule out our results being driven by spurious correlations

or other alternative factors for which we did not account.

This paper fills a gap in the literature and explores some of our previous work on the role migrants play in facilitating the transfer of knowledge across borders (Bahar and Rapoport, 2018; Choudhury, 2016; Choudhury and Kim, 2018). Specifically, it examines a particular channel through which inventor migrants–a small and very particular subset of high-skilled migrants–can heavily influence innovation dynamics in their receiving countries.

By providing robust results of how migration affects the transfer of specific technologies across borders, from the home country of the migrant to the host country, using a large number of countries, studied over two decade-long periods, our study contributes to the literature on migrants and innovation (e.g., Kerr, 2008; Agrawal et al., 2008; Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010; Freeman and Huang, 2015; Ganguli, 2015; Bosetti et al., 2015; Choudhury, 2016; Akcigit et al., 2017; Breschi et al., 2017; Bernstein et al., 2018; Miguélez, 2018; Choudhury and Kim, 2018; Doran and Yoon, 2019). More broadly, our findings indicate that migrant inventors can play an important role in shaping the patent production function in their host countries. Arguably, these dynamics driven by migrant inventors play an important role in improving other economic outcomes that follow patenting and innovation, such as productivity and, ultimately, economic growth. Hence, this study is another piece of evidence that the overall medium- to long-term economic gains from migration are large and persistent over time.

References

Agrawal, Ajay, Devesh Kapur, and John McHale. "How do spatial and social proximity influence knowledge flows? Evidence from patent data." *Journal of Urban Economics* 64, 2: (2008) 258–269.

Agrawal, Ajay, Devesh Kapur, John McHale, and Alexander Oettl. "Brain drain or brain bank? The impact of skilled emigration on poor-country innovation." *Journal of Urban Economics* 69, 1: (2011) 43–55.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas. "Immigration and the Rise of American Ingenuity." *American Economic Review* 107, 5: (2017) 327–331.

Almeida, Paul, Anupama Phene, and Sali Li. "The Influence of Ethnic Community Knowledge on Indian Inventor Innovativeness." *Organization Science* , May 2019: (2014) 141223041331,004.

Bahar, Dany, Ricardo Hausmann, and Cesar A. Hidalgo. "Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion?" *Journal of International Economics* 92, 1: (2014) 111–123.

Bahar, Dany, and Hillel Rapoport. "Migration, Knowledge Diffusion and the Comparative Advantage of Nations." *The Economic Journal* 128, 612: (2018) F273–F305.

Bahar, Dany, Samuel Rosenow, Ernesto Stein, and Rodrigo Wagner. "Export take-offs and acceleration: Unpacking cross-sector linkages in the evolution of comparative advantage." *World Development* 117: (2019) 48–60.

Balassa, B. "Trade Liberalisation and Revealed Comparative Advantage." *The Manchester School* 33, 2: (1965) 99–123.

Bernstein, Shai, Rebecca Diamond, Timothy James McQuade, and Beatriz Pousada. "The Contribution of High-Skilled Immigrants to Innovation

in the United States." *Stanford University Graduate School of Business Working Paper* .

Berry, Heather. "The Influence of Multiple Knowledge Networks on Innovation in Foreign Operations." *Organization Science* 29, 5: (2018) 855–872.

Bosetti, Valentina, Cristina Cattaneo, and Elena Verdolini. "Migration of skilled workers and innovation: A European Perspective." *Journal of International Economics* 96, 2: (2015) 311–322. http://dx.doi.org/10.1016/j.jinteco.2015.04.002.

Breschi, Stefano, and Francesco Lissoni. "Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows." *Journal of Economic Geography* 9, 4: (2009) 439–468.

Breschi, Stefano, Francesco Lissoni, and Franco Malerba. "Knowledge-relatedness in firm technological diversification." *Research Policy* 32, 1: (2003) 69–87.

Breschi, Stefano, Francesco Lissoni, and Ernest Miguelez. "Foreign-origin inventors in the USA: Testing for diaspora and brain gain effects." *Journal of Economic Geography* 17, 5: (2017) 1009–1038.

Burchardi, Konrad B., Thomas Chaney, and Tarek A. Hassan. "Migrants, Ancestors, and Foreign Investments." *The Review of Economic Studies* , June.

Burchardi, Konrad B., Thomas Chaney, Tarek A. Hassan, Lisa Tarquinio, and Stephen J. Terry. "Immigration, Innovation and Growth." *Working Paper* .

Card, David. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19, 1: (2001) 22–64.

Choudhury, Prithwiraj. "Return migration and geography of innovation in MNEs: a natural experiment of knowledge production by local workers

reporting to return migrants." *Journal of Economic Geography* 16, 3: (2016) 585–610.

Choudhury, Prithwiraj, and Do Yoon Kim. "The Ethnic Migrant Inventor Effect: Codification And Recombination of Knowledge Across Borders.", 2018.

Cipolla, Carlo M. *Before the Industrial Revolution: European Society and Economy 1000-1700*. London: Methuen, 1976, 1st edition.

Criscuolo, Paola. "The 'home advantage' effect and patent families. A comparison of OECD triadic patents, the USPTO and the EPO." *Scientometrics* 66, 1: (2005) 23–41.

De Noni, Ivan, Luigi Orsi, and Fiorenza Belussi. "The role of collaborative networks in supporting the innovation performances of lagging-behind European regions." *Research Policy* 47, 1: (2018) 1–13. `https://doi.org/10.1016/j.respol.2017.09.006`.

Doran, Kirk, and Chungeun Yoon. "Immigration and Invention: Evidence from the Quota Acts." *mimeo* .

Foley, C Fritz, and William R Kerr. "Ethnic Innovation and U.S. Multinational Firm Activity." *Management Science* 59, 7: (2013) 1529–1544.

Freeman, Richard B., and Wei Huang. "Collaborating with People Like Me: Ethnic Co-Authorship within the U.S." *Journal of Labor Economics* 33, S1: (2015) S289–S318.

Ganguli, Ina. "Immigration and Ideas: What Did Russian Scientists "Bring" to the United States?" *Journal of Labor Economics* 33, S1: (2015) S257–S288.

Glismann, Hans H., and Ernst-Jürgen Horn. "Comparative Invention Performance of Major Industrial Countries: Patterns and Explanations." *Management Science* 34, 10: (1988) 1169–1187.

Glitz, Albrecht, and Erik Meyersson. "Industrial Espionage and Productivity." *IZA Discussion Paper Series* , 10816: (2017) 1–50.

Hausmann, Ricardo, César A Hidalgo, Sebastián Bustos, Michele Coscia, Alexander Simoes, and Muhammed A. Yildirim. *The Atlas of Economic Complexity: Mapping Paths to Prosperity.* Cambridge, MA: MIT Press, 2014.

Henderson, Rebecca, Adam Jaffe, Manuel Trajtenberg, Peter Thompson, and Melanie Fox-Kean. "Patent citations and the geography of knowledge spillovers: A reassessment: Comment." *American Economic Review* 95, 1: (2005) 461–466.

Horrace, William C., and Ronald L. Oaxaca. "Results on the bias and inconsistency of ordinary least squares for the linear probability model." *Economics Letters* .

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics* 2, 2: (2010) 31–56.

Jaffe, Adam B. "Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value." *The American Economic Review* 76, 5: (1986) 984–1001.

Jaffe, Adam B. "Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers." *Research Policy* 18, 2: (1989) 87–97.

Kerr, Sari Pekkala, and William R. Kerr. "Global Collaborative Patents." *The Economic Journal* 128, 612: (2018) F235–F272.

Kerr, W R. "Ethnic Scientific Communities and International Technology Diffusion." *Review of Economics and Statistics* 90, 3: (2008) 518–537.

Kerr, W.R., and W. Lincoln. "The supply side of innovation: H-1B visa reforms and US ethnic invention." *Journal of Labor Economics* 28, 3.

Lissoni, Francesco. "International migration and innovation diffusion: an eclectic survey." *Regional Studies* 52, 5: (2018) 702–714.

MacKinnon, JG, and Lonnie Magee. "Transforming the Dependent Variable in Regression Models." *International Economic Review* 31, 2: (1990) 315–339.

Miguélez, Ernest. "Inventor diasporas and the internationalization of technology." *World Bank Economic Review* 32, 1: (2018) 41–63.

Miguelez, Ernest, and Carsten Fink. "Measuring the international mobility of inventors: A new database." *The International Mobility of Talent and Innovation: New Evidence and Policy Implications* , 8: (2017) 114–161.

Miguelez, Ernest, and Claudia Noumedem Temgoua. "Inventor migration and knowledge flows: A two-way communication channel?" *Research Policy* 103914.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. "German Jewish Émigrés and US Invention." *American Economic Review* 104, 10: (2014) 3222–3255.

OECD. "Revealed technology advantage in selected fields." .

———. "Patents by main technology and by International Patent Classification (IPC)." `https://www.oecd-ilibrary.org/content/data/data-00508-en`.

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

Singer, Judith D., and John B. Willett. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford: Oxford University Press, 2009.

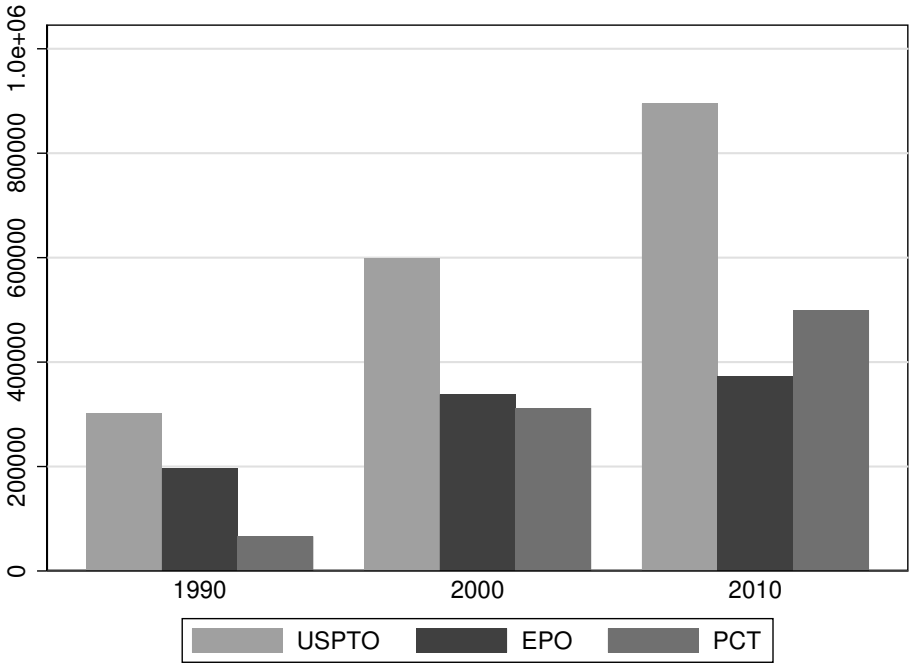
Singh, Jasjit, and Matt Marx. "Geographic Constraints on Knowledge Spillovers: Political Borders vs. Spatial Proximity." *Management Science* 59, 9: (2013) 2056–2078.

Soete, Luc. "The impact of technological innovation on international trade patterns: The evidence reconsidered." *Research Policy* 16, 2-4: (1987) 101–130.

Stock, JH, and M Yogo. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models*, edited by DWK Andrews, New York: Cambridge University Press, 2005, 80–108.

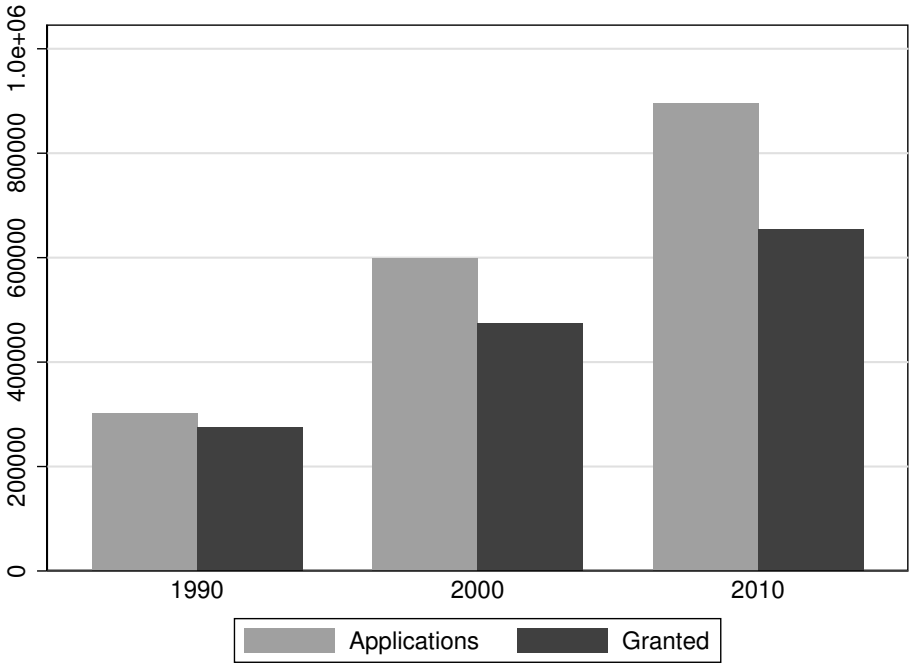
Thompson, Peter, and Melanie Fox-Kean. "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment." *The American Economic Review* 95, 1: (2005) 450–460.

Figure 1: Patent applications by year (USPTO, EPO, and PCT)



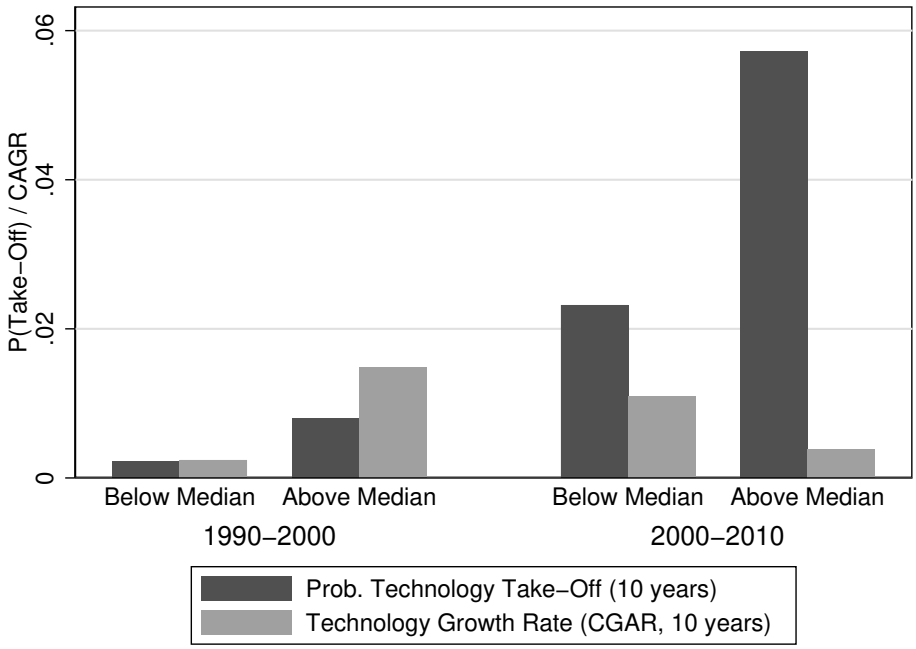
This figure presents the total number of patent applications to the USPTO, EPO, and PCT in years 1990, 2000, and 2010 around the globe.

Figure 2: Patent applications and granted (USPTO), by year



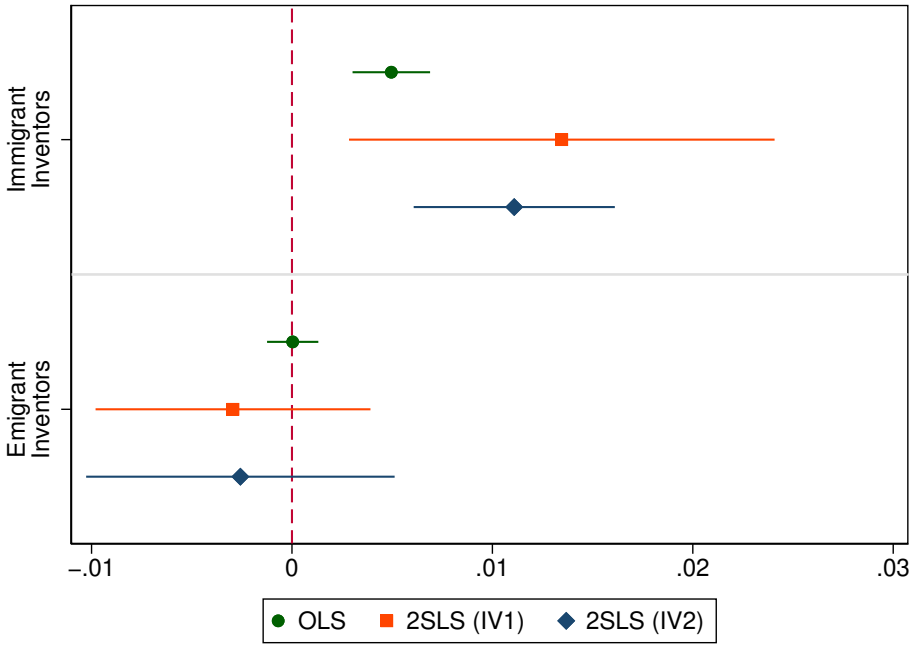
This figure presents the total number of patent applications and patents granted in years 1990, 2000, and 2010 around the globe, based on the records of the USPTO.

Figure 3: Patent application take-offs, raw data



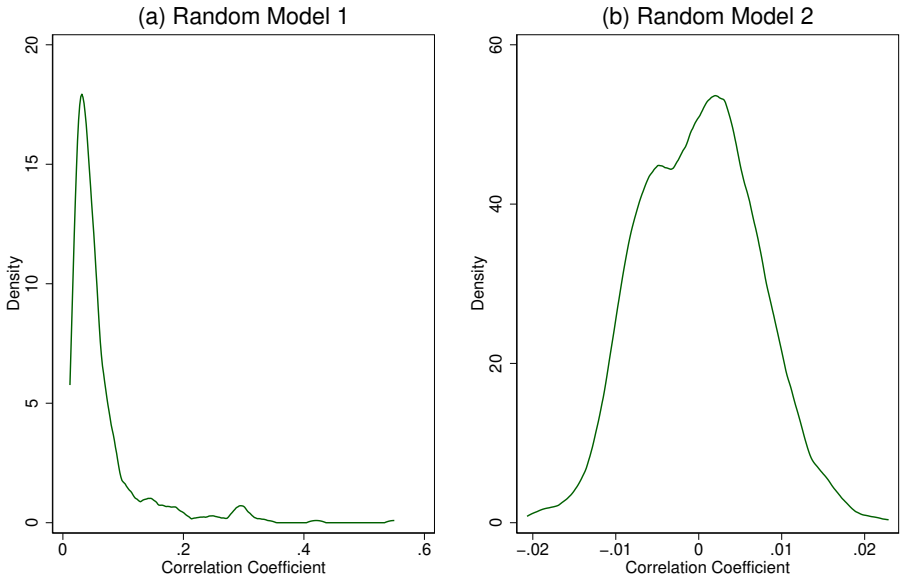
This figure presents the average probability of a patent technology take-off and the average CAGR (using patent applications) for country-technology pairs with a stock of immigrant inventors from countries that specialize in that same technology (e.g., file patents in that technology subclass with an RTA above 1) below and above the sample median. The stock of immigrants is scaled by population size of the receiving country. The figure is based on simple averages, with no controls whatsoever.

Figure 4: OLS vs. 2SLS estimators (Technology Take-offs)



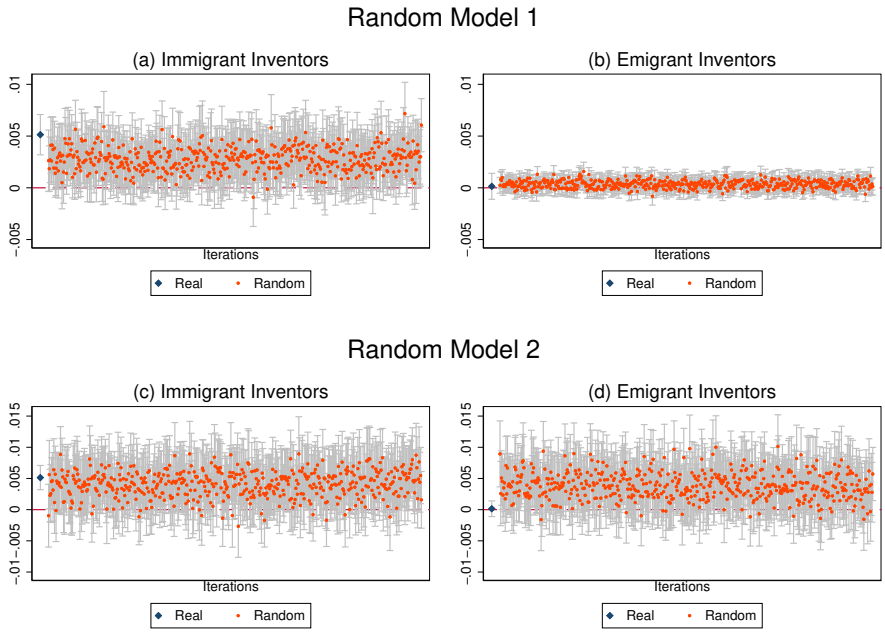
This figure plots the point estimates and their corresponding 95% confidence intervals (represented by whiskers) of the OLS and two different 2SLS estimations for both β^{IM} and β^{IM}, based on the results presented in Panel A of Table 2.

Figure 5: Correlations between real and random-inventor migrants



This figure plots the kernel distributions between real and random inventors based on 500 iterations. The left panel is based on Random Model 1, which generates a random number maintaining the original distributional characteristics of the actual variable; the right panel is based on Random Model 2, which generates a random vector of migrant inventors, from 0 to 1, without any data restrictions whatsoever.

Figure 6: Summary of 500 estimations using random inventor figures (OLS)



This figure plots the estimators of β^{IM} (left panel) and β^{EM} (right panel) when substituting the real number of migrant inventors between countries with a random one, for each of 500 iterations. The results are based on an OLS estimation that includes both regressors simultaneously. The upper row is based on a randomization approach such that the real and the random number of inventors have the same sample mean and distribution. The lower panel is based on a randomization approach that replaces the actual number of inventors with a random number, with no restrictions whatsoever, distributed uniformly from 0 to 1. The figure also includes, for reference, the estimators using the actual number of migrant inventors, reported in our main results, and marked with a diamond-shaped symbol. Whiskers represent 95% confidence intervals, based on SE clustered at the country level.

Table 1: Summary Statistics

Variable	N	Mean	sd	Min	Max
Panel A - Take-off sample ($RTA_{c,p,t_o} = 0$)					
Take-off technology (patent applications)	105,304	0.022	0.15	0.0	1.0
Immigrant inventors (weighted)	105,304	24.245	135.54	0.0	8,863.0
Emigrant inventors (weighted)	105,304	79.775	411.80	0.0	7,624.0
Total FDI (billion USD, weighted)	105,304	32.504	216.92	0.0	9,966.9
Total trade (billion USD, weighted)	105,304	106.827	258.35	0.0	4,835.1
Total immigrants lagged IV (thous.,weighted)	105,304	69.439	328.88	0.0	8,701.1
Total emigrants lagged IV (thous.,weighted)	105,304	141.678	486.31	0.0	8,753.7
Immigrant inv. push-pull IV (weighted)	105,304	17.727	91.16	0.0	6,871.3
Emigrant inv. push-pull IV (weighted)	105,304	55.170	214.60	0.0	4,387.0
Panel B - Growth sample ($patents_{c,p,t_o} > 0$)					
CAGR technology (patent applications)	18,386	0.007	0.08	-0.3	0.8
Baseline patent apps	18,386	16.447	119.13	1.0	9,730.0
Immigrant inventors (weighted)	18,386	709.735	2,116.58	0.0	26,642.0
Emigrant inventors (weighted)	18,386	610.600	1,138.10	0.0	8,229.0
Total FDI (billion USD, weighted)	18,386	836.282	1,567.10	0.0	14,487.1
Total trade (billion USD, weighted)	18,386	1,109.598	1,379.50	0.0	12,273.9
Total immigrants lagged IV (thous.,weighted)	18,386	499.439	903.89	0.0	8,708.0
Total emigrants lagged IV (thous.,weighted)	18,386	448.861	705.99	0.0	8,699.5
Immigrant inv. push-pull IV (weighted)	18,386	468.506	1,150.29	0.0	12,917.9
Emigrant inv. push-pull IV (weighted)	18,386	402.796	674.82	0.0	4,975.3

This table presents the sample summary statistics for the variables used in the paper. The upper panel presents the sample used in the estimations of technology take-offs, where we limit the sample to those country-technology observations that have no patents granted in the beginning of the 1990-2000 and 2000-2010 periods). The lower panel presents results used in the estimations of patent growth regressions, where we limit our observations to those country-technology pairs with number of patents above zero at the beginning of the 1990-2000 and 2000-2010 periods.

Table 2: Migrant inventors and patent applications take-offs and growth

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0050		0.0050	0.0111		0.0135	0.0098		0.0111
	(0.001)***		(0.001)***	(0.004)***		(0.005)**	(0.002)***		(0.003)***
Emigrant inventors		0.0017	0.0000		0.0059	-0.0029		0.0098	-0.0026
		(0.001)***	(0.001)		(0.003)**	(0.003)		(0.002)***	(0.004)
Total FDI	-0.0001	-0.0001	-0.0001	-0.0003	-0.0002	-0.0002	-0.0002	-0.0004	-0.0002
	(0.000)	(0.000)	(0.000)	(0.000)**	(0.000)*	(0.000)	(0.000)***	(0.000)***	(0.000)
Total Trade	0.0002	0.0003	0.0002	-0.0003	-0.0002	-0.0001	-0.0002	-0.0007	-0.0000
	(0.000)	(0.000)**	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)*	(0.000)**	(0.000)
N	105304	105304	105304	105304	105304	105304	105304	105304	105304
r2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
KP F Stat				60.59	37.42	40.91	754.69	99.14	26.79
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0045		0.0034	0.0243		0.0092	-0.0203		-0.0070
	(0.001)***		(0.001)***	(0.010)**		(0.010)	(0.048)		(0.054)
Emigrant inventors		0.0042	0.0029		0.0250	0.0198		0.0253	0.0241
		(0.001)***	(0.001)**		(0.016)	(0.021)		(0.029)	(0.034)
Total FDI	0.0003	0.0003	0.0003	0.0002	0.0000	0.0000	0.0006	0.0000	0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Total Trade	0.0027	0.0021	0.0016	-0.0039	-0.0081	-0.0086	0.0109	-0.0082	-0.0053
	(0.002)	(0.001)	(0.001)	(0.003)	(0.008)	(0.008)	(0.018)	(0.014)	(0.032)
Baseline patent apps, log	-0.0178	-0.0176	-0.0178	-0.0186	-0.0173	-0.0178	-0.0169	-0.0173	-0.0171
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Previous Exports Growth	-0.1423	-0.1436	-0.1424	-0.1351	-0.1419	-0.1390	-0.1513	-0.1419	-0.1446
	(0.039)***	(0.039)***	(0.039)***	(0.041)***	(0.040)***	(0.038)***	(0.042)***	(0.039)***	(0.047)***
Zero Exports in t-1	-0.0015	-0.0013	-0.0014	-0.0016	-0.0005	-0.0008	-0.0014	-0.0005	-0.0005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
N	18349	18349	18349	18349	18349	18349	18349	18349	18349
r2	0.55	0.55	0.55	0.53	0.53	0.53	0.52	0.53	0.53
KP F Stat				8.53	5.83	7.59	0.40	2.50	0.25

This table presents results of the estimation of Specification (1). Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair was zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

Table 3: Migrant inventors and patent applications, falsification test 1

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	-0.0054		-0.0048	-0.0057		-0.0018	-0.0112		-0.0117
	(0.002)**		(0.002)**	(0.009)		(0.013)	(0.006)*		(0.007)*
Emigrant inventors		-0.0028	-0.0022		-0.0050	-0.0040		-0.0018	0.0011
		(0.001)**	(0.001)		(0.007)	(0.010)		(0.004)	(0.004)
Total FDI	0.0003	0.0004	0.0004	0.0003	0.0004	0.0004	0.0003	0.0004	0.0003
	(0.000)**	(0.000)***	(0.000)**	(0.000)**	(0.000)**	(0.000)*	(0.000)**	(0.000)**	(0.000)**
Total Trade	-0.0007	-0.0008	0.0016	-0.0005	0.0018	0.0017	0.0030	-0.0020	0.0021
	(0.003)	(0.004)	(0.003)	(0.007)	(0.009)	(0.009)	(0.004)	(0.005)	(0.005)
N	105304	105304	105304	105304	105304	105304	105304	105304	105304
r2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
KP F Stat				6.56	5.52	5.59	110.58	66.37	101.29
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	-0.0084		-0.0068	-0.0169		-0.0272	-0.1088		-0.0537
	(0.002)***		(0.002)***	(0.009)*		(0.011)**	(0.024)***		(0.108)
Emigrant inventors		-0.0065	-0.0056		-0.0027	0.0164		-0.0455	-0.0242
		(0.002)***	(0.002)***		(0.010)	(0.013)		(0.015)***	(0.048)
Total FDI	0.0001	0.0001	0.0001	0.0001	0.0000	0.0000	0.0007	0.0004	0.0006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total Trade	-0.0047	-0.0041	-0.0001	0.0013	-0.0079	-0.0080	0.0669	0.0357	0.0524
	(0.005)	(0.005)	(0.005)	(0.010)	(0.014)	(0.013)	(0.021)***	(0.022)	(0.030)*
Baseline patent apps, log	-0.0180	-0.0183	-0.0184	-0.0182	-0.0180	-0.0172	-0.0205	-0.0213	-0.0210
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Previous Exports Growth	-0.1491	-0.1464	-0.1489	-0.1523	-0.1461	-0.1553	-0.1869	-0.1485	-0.1675
	(0.038)***	(0.040)***	(0.038)***	(0.034)***	(0.040)***	(0.033)***	(0.027)***	(0.037)***	(0.056)***
Zero Exports in t-1	-0.0008	-0.0012	-0.0008	-0.0004	-0.0012	-0.0002	0.0046	-0.0003	0.0021
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.007)
N	18349	18349	18349	18349	18349	18349	18349	18349	18349
r2	0.55	0.55	0.55	0.55	0.55	0.52	0.32	0.48	0.47
KP F Stat				42.33	12.48	7.61	9.98	13.54	0.23

This table presents results of the estimation of Specification (1), with a slight modification. Instead of including in the RHS the number of migrant inventors from and in countries that specialize in each technology in the baseline year, we use the number of migrant inventors from and in countries with *zero patent applications in each technology at the beginning of the period. Columns 1-3 show OLS estimations, while columns 4-6 show results for 2SLS regressions. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair were zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.*

$^{*}p < 0.10,^{**}\ p < 0.05,^{***}\ p < 0.01$

Table 4: Migrant inventors and patent applications, subsample analysis

		OLS		**2SLS (IV1)**		**2SLS (IV2)**	
	N	β_{im}	β_{em}	β_{im}	β_{em}	β_{im}	β_{em}
All Observations	105304	0.005***	-0.000	0.013**	-0.004	0.011***	-0.003
Non OECD	76695	0.006***	0.000	0.015**	-0.004	0.021***	-0.001
OECD	28609	0.003	0.000	0.015**	-0.003	0.007**	-0.005
Period 1990-2000	54042	0.000	0.000	0.007	-0.003	0.002*	-0.000
Period 2000-2010	51262	0.007***	-0.000	0.014**	-0.004	0.017***	-0.005
Chemistry; Metallurgy	13564	0.009***	0.001	0.016	-0.005	0.012**	0.001
Electricity	7571	0.005	0.002	0.008	0.000	0.017***	-0.003
Fixed Constructions	4985	-0.000	0.002	0.016	-0.010	0.005	0.003
Human Necessities	13211	0.009***	-0.001	0.012	-0.006	0.031***	-0.024
Mechanical Engineering	15958	0.004	-0.002	0.018	-0.005	0.019**	-0.008
Performing Operations; Transporting	27529	0.001	-0.003	0.011	-0.006	0.007	-0.014
Physics	12788	0.006*	0.001	0.023	-0.005	0.016***	-0.004
Textiles; Paper	6663	0.003*	-0.001	0.010	-0.006	0.012***	-0.006

This table summarizes OLS regressions for different cuts of the sample. The reported beta coefficients are standardized to have zero mean and unit standard deviation, for comparison purposes. The dependent variable in all specifications is technology take-off. Significance levels reported based on SE clustered at the country level

Online Appendix for

Migration inventors and the technological advantage of nations

Dany Bahar, Prithwiraj Choudhury, and Hillel Rapoport

February 17, 2020

A Extended summary statistics

Table A1 presents summary statistics for each of the 95 countries in the sample. The table presents the average probability of take-off for each country (presented in percentage format) based on observations of country-technology pairs with no patent applications in the baseline year for each period. Similarly, the growth rate (also presented in percentage format) is based on observations of country-technology pairs with more than zero patent applications in the baseline period. The * symbol next to a country name implies that the country is part of the OECD prior to the first period studied (e.g., it does not count countries as OECD members if they became so during the 1990s or 2000s).

[Table A1 about here.]

Table A1: Summary statistics by country

	1990-2000		2000-2010			1990-2000		2000-2010	
Country	**Take off**	**CAGR**	**Take Off**	**CAGR**	**Country**	**Take off**	**CAGR**	**Take Off**	**CAGR**
United Arab Emirates	0.00	-3.99	5.54	-5.88	South Korea	1.39	10.24	5.58	6.77
Argentina	0.32	-3.46	7.03	-3.50	Kuwait	0.00	-6.70	6.45	-3.33
Armenia	0.00	.	2.01	-4.34	Lebanon	0.00	-6.70	1.71	-4.13
Australia *	0.82	1.10	6.92	0.92	Liechtenstein	0.32	-3.72	2.12	-5.34
Austria *	0.51	-0.04	7.90	0.41	Sri Lanka	0.00	-6.09	1.56	-5.86
Belgium *	1.16	0.82	7.84	0.54	Lithuania	0.00	.	2.79	-2.66
Bulgaria	0.00	-6.16	3.13	-2.47	Luxembourg *	0.49	-4.64	5.53	-3.67
Bosnia and Herzegovina	0.00	.	1.08	-6.70	Latvia	0.00	.	2.32	0.84
Belarus	0.15	.	2.02	-2.96	Morocco	0.00	-6.70	1.10	-5.58
Bermuda	0.00	-6.70	0.47	-7.11	Monaco	0.16	-4.69	1.60	-5.78
Brazil	0.84	-1.76	10.36	0.62	Moldova	0.00	-6.70	1.69	-6.70
Canada *	0.55	3.17	10.14	0.94	Mexico	0.85	-3.15	10.17	-2.80
Switzerland *	1.30	0.00	2.16	0.21	Macedonia, FYR	0.00	.	0.00	-6.70
Chile	0.31	-5.68	6.55	-2.70	Malta	0.00	0.00	1.54	-7.93
China	2.16	2.36	15.42	15.21	Mongolia	0.00	-6.70	0.31	.
Colombia	0.00	-3.15	4.91	-3.42	Malaysia	0.47	0.30	9.79	-0.13
Costa Rica	0.00	-7.31	1.40	-4.29	Nigeria	0.00	-6.70	0.15	-2.04
Cuba	0.31	.	0.31	-3.44	Netherlands *	1.29	0.57	8.37	1.08
Cayman Islands	0.00	-6.70	2.32	-5.02	Norway *	1.14	-0.94	7.18	-1.12
Cyprus	0.00	-6.70	3.56	-3.92	New Zealand *	0.52	-1.44	6.02	-1.23
Czech Republic	0.48	-4.77	8.96	-0.94	Pakistan	0.00	-6.70	3.10	-6.70
Germany *	1.02	3.44	1.25	0.40	Panama	0.00	-6.70	0.93	-6.70
Denmark *	1.30	0.06	6.07	-0.59	Peru	0.00	-7.59	2.79	-6.70
Algeria	0.00	-6.70	1.08	-6.70	Philippines	0.79	-5.35	5.56	-4.13
Ecuador	0.00	-6.70	0.62	-6.70	Poland	0.16	-3.75	10.07	-1.36
Egypt	0.00	-6.70	5.38	-4.09	North Korea	0.00	-6.70	0.46	-6.70
Spain *	1.76	-0.21	10.44	1.37	Portugal *	0.47	-4.36	6.48	-1.29
Estonia	0.00	.	5.15	-0.62	Russian Federation	2.39	-2.36	8.11	-1.76
Finland *	1.89	1.24	5.52	-1.37	Saudi Arabia	0.31	-1.99	12.97	3.97
France *	0.00	0.77	5.96	0.81	Singapore	1.91	4.87	7.64	0.15
United Kingdom *	0.00	1.51	3.40	0.63	El Salvador	0.00	.	0.00	-6.70
Georgia	0.00	.	1.54	-6.70	Slovakia	0.77	.	4.18	-4.48
Greece *	0.31	-5.60	6.08	-1.61	Slovenia	0.92	.	3.91	-2.37
Guatemala	0.00	-6.70	0.46	-7.93	Sweden *	1.35	2.73	4.74	-1.63
Hong Kong	1.36	1.60	7.35	-0.95	Seychelles	0.00	-6.70	0.77	.
Croatia	0.61	.	4.11	-3.53	Thailand	0.47	-1.90	6.81	-4.04
Hungary	0.00	-4.43	7.00	-1.97	Trinidad and Tobago	0.00	-6.70	2.48	-7.23
Indonesia	0.00	-4.02	3.50	-3.80	Tunisia	0.00	-6.70	1.55	-2.23
India	1.15	4.39	14.42	5.68	Turkey *	0.00	-5.21	8.32	-1.16
Ireland *	0.35	-0.74	6.69	0.18	Ukraine	0.77	.	5.52	-3.63
Iran	0.00	-6.70	8.83	-4.14	Uruguay	0.00	-6.70	1.39	-6.89
Iceland *	0.77	-4.46	3.06	-3.39	United States *	0.00	3.05	3.03	-0.04
Israel	1.62	3.15	3.10	0.72	Uzbekistan	0.00	.	0.15	-6.70
Italy *	0.79	0.98	7.01	0.21	Venezuela	0.00	-3.86	1.17	-5.12
Jamaica	0.00	-6.70	0.93	-6.70	South Africa	0.00	-3.26	5.27	-3.31
Jordan	0.00	-6.70	2.15	-6.70	Zimbabwe	0.00	.	0.15	-6.70
Japan *	0.00	2.72	4.48	0.37		.	.	.	.
Kazakhstan	0.00	.	1.24	-6.70		.	.	.	.
Kenya	0.00	-4.46	0.78	-3.26		.	.	.	.

The table presents values for the probability of take-off and the growth rate (CAGR), both in percentage terms, averaged across all technology subclasses for each country in each period in our sample. Consistent with our analysis, the table presents the average probability of take-off for each country based on observations of country-technology pairs with no patent applications in the baseline year for each period. Similarly, the growth rate is based on observations of country-technology pairs with more than zero patent applications in the baseline period. The * symbol indicates whether the country is an OECD member (based on the 1980 cutoff, prior to the first period of analysis in our exercise).

B Explanatory power of bilateral migration on future bilateral inventor stocks

As noted in the paper, one plausible concern about our identification strategy of instrumenting current stocks of migrant inventors with our instruments is that the first-stage correlation of the 2SLS estimation is artificially driven up by the weighting scheme we use in our baseline specification. However, this is not the case, as we show in the following exercise; we estimate the following specification using data for years 1990 and 2000:

$$inventors_{c,c',t} = \beta_{IV} IV_{c,c',t} + \gamma_c + \gamma_{c'} + \theta_t + \varepsilon_{c,c',t}$$

Where c and c' are countries; t is year; $inventors_{c,c',t}$ is the stock of migrant inventors from country c' in country c at time t; and $IV_{c,c',t}$ is one of the two IVs we used for our identification strategy: the stock of total migrants from country c' in country c at time t-30 (that is, for years 1960 and 1970), as well as the predicted stock of inventors using push-pull factors. Both terms in the regression are transformed using the inverse hyperbolic sine and, thus, β_{IV} can be interpreted as an elasticity. γ_c and $\gamma_{c'}$ are receiving and sending country fixed effects, respectively; θ_t represents year fixed effects. The last term represents the error.

The estimation for this specification is presented in Table B1. The table reports the estimation of β_{IV} for each instrument using three different estimations that vary with the inclusion of fixed effects, as well as standard errors in parenthesis and t-stats below them for the purpose of this exercise. Columns 1 and 4 include only year fixed effects; Columns 2 and 5 include year fixed effects as well as sending country and receiving country fixed effects; Columns 3 and 6 include sending country and year (combined) fixed effects as well as receiving country-by-year fixed effects. Columns 1 to 3 use our first instrumental variable, the lagged stock of migrants; Columns 4 to 6 use the predicted number of inventor migrants. In all columns, we see that there is a positive elasticity and, perhaps more importantly, a strong explanatory power with t-stats. Thus, we conclude that the success of our first stage is

not driven by mechanics due to the transformation of the variables.

[Table B1 about here.]

Table B1: Explanatory power of lagged migration on inventor migrants

Dependent variable: Stock of Immigrant Inventors						
	est1	est2	est3	est4	est5	est6
Immigrant stock (t-30)	0.0754	0.0362	0.0356			
	(0.016)***	(0.007)***	(0.007)***			
	4.701	5.168	5.126			
Immigrant inventors (Push-Pull prediction)				0.8446	0.8363	0.8374
				(0.044)***	(0.038)***	(0.036)***
				19.024	22.053	23.028
Constant	-0.0704	0.0254	0.0268	-0.0100	-0.0088	-0.0089
	(0.020)***	(0.017)	(0.017)	(0.002)***	(0.006)	(0.006)
	-3.451	1.476	1.573	-4.019	-1.498	-1.583
N	42924	42924	42924	42924	42924	42924
Adj. R2	0.16	0.35	0.41	0.72	0.74	0.74
θ_t	Y	Y	-	Y	Y	-
$\gamma_c, \gamma_{c'}$	N	Y	-	N	Y	-
$\gamma_c \times \theta_t, \gamma_{c'} \times \theta_t$	-	-	Y	-	-	Y

This table estimates the elasticity of the stock of migrant inventors to the stock of total migrants 30 years before, using bilateral figures. Both left-hand side and right-hand side variables are transformed using the inverse hyperbolic sine (asinh). Columns 1 and 4 include only year fixed effects; columns 2 and 5 include receiving and sending country fixed effects, as well as year fixed effects; columns 3 and 6 include receiving country and year (combined) as well as sending country and year (combined) fixed effects. The estimation uses years 1990 and 2000 as baseline years. Standard errors are clustered at the receiving country and sending country levels and are presented in parenthesis. t-stats are presented below the standard errors

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

C Alternative methods for dependent binary variables

As noted, our dependent variable measuring technology take-offs is a binary variable. Hence, our OLS estimation represents a linear probability model (LPM). We believe that given the computational difficulties posed by estimating nonlinear models using high-dimensional fixed effects, LPM is a reasonable choice. However, for robustness purposes, we reestimate our main specification using the complementary log-log estimator. This particular maximum likelihood estimator, used to estimate models where the dependent variable responds to a binary distribution, is a more advisable option than logit or probit if the probability of take-off is small, as in our case (Singer and Willett, 2009). For computational reasons, to limit the number of simultaneous fixed effects, we estimate this c-log-log model for the 2000 to 2010 decade only, focusing only on take-offs, naturally. Table C1 presents the results, which are robust to the ones presented in the main body of the paper.

[Table C1 about here.]

Alternatively, we also consider the approach of Horrace and Oaxaca (2006) for binary data. The approach suggests, after the first estimation of a linear probability model, dropping from the sample the observations for which the predicted value falls out of the unit interval and using this subsample to reestimate the linear probability model. Horrace and Oaxaca (2006) show that this approach may reduce the potential biases of the linear probability models. Table C2 reports the results using this method. In fact, very few observations in our sample have a predicted value outside of the 0 to 1 range (less than half a percent), so the number of observations in this estimation is very close to those in Panel A of Table 2, therefore arriving at almost identical results. This suggests the potential bias from linear probability models is minimal, if anything.

[Table C2 about here.]

Table C1: Migrant inventors and patent applications take-offs, c-log-log estimation

Dependent variable: Patent applications take-off (binary)			
	est1	est2	est3
Take off patent class			
Immigrant inventors	0.0856		0.1001
	(0.036)**		(0.037)***
Emigrant inventors		-0.0221	-0.0580
		(0.045)	(0.047)
Total FDI	0.0056	0.0067	0.0069
	(0.010)	(0.010)	(0.010)
Total Trade	0.0220	0.0461	0.0342
	(0.023)	(0.023)**	(0.024)
N	41221	41221	41221
r2_p			

This table presents results of the estimation of Specification (1), focusing on take-offs (a binary dependent variable), using a (maximum likelihood) c-log-log estimator. All specifications include country and technology fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**}\ p < 0.05,^{***}\ p < 0.01$

Table C2: Migrant inventors and patent applications take-offs, HO (2006)

Dependent variable: Patent applications take-off (binary)			
	est1	est2	est3
Immigrant inventors	0.0050		0.0050
	(0.001)***		(0.001)***
Emigrant inventors		0.0017	0.0000
		(0.001)***	(0.001)
Total FDI	-0.0001	-0.0001	-0.0001
	(0.000)	(0.000)	(0.000)
Total Trade	0.0002	0.0003	0.0002
	(0.000)	(0.000)**	(0.000)
N	104792	104792	104792
r2	0.09	0.09	0.09

This table presents results of the estimation of Specification (1), focusing on take-offs (a binary dependent variable), using the methodology suggested by Horrace and Oaxaca (2016). All specifications include country and technology fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

D Alternative dependent variables

For robustness purposes, we extend our exercise to include alternative take-off and growth measures as our dependent variable.

Table D1 replicates the estimation shown in Panel A of Table 2, this time using a modified version of take-offs that does not rely on RTA thresholds. The measure uses a much simpler way to account for new technology subclasses in a given country-technology pair based on zero or larger than zero patent applications. In this case, the dependent variable is measured as:

$$Y_{c,p,t\to T} = 1 \text{ if } patents_{c,p,t} = 0 \text{ and } patents_{c,p,T} \geq 1 \ ,$$

where $T - t = 10$.

[Table D1 about here.]

The results in Table D1 are qualitatively consistent with those in Panel A of Table 2, with larger point estimates, as is expected. The 2SLS results using the first set of instruments is much noisier (thus, we are unable to find results that are statistically significant), whereas the estimations using the second set of instrument show results consistent with our main estimation. Note that when using these results (instead of the RTA-based measure as our dependent variable), we are unable to conclude anything about gaining technological advantage in certain technology subclasses due to inventor migrants.

We also replicate our main specification using an alternative measure of growth that includes those observations that start off with zero patents. This measure is the symmetric percentage change and is defined as:

$$Y_{c,p,t\to T} = \frac{patents_{c,p,T} - patents_{c,p,t}}{0.5 * (patents_{c,p,T} + patents_{c,p,t})} \times \frac{1}{T-t}$$

where $T - t = 10$. In a sense, this growth measure allows us to include all observations, as opposed to CAGR, since we can include those country-technology pairs for which $patents_{c,p,t} = 0$ as their symmetric percentage change (SPC) is defined. This measure also presents a symmetric growth rate so that a change from $patents_{c,p,t} = x_1$ to $patents_{c,p,T} = x_2$ represents

the same percentage change (in absolute value) than from $patents_{c,p,T} = x_2$ to $patents_{c,p,t} = x_1$. Table D2 presents the results using all the sample (i.e., including those country-technology pairs that start off with zero patent applications).

[Table D2 about here.]

The OLS results of Table D2 show that when using this alternative dependent variable, we find results consistent with our main estimations: Immigrant inventors explain faster growth rates in patent applications in the same specialized technologies of their origin countries. The 2SLS results are somewhat weaker in terms of statistical significance for the first set of instruments, but not for the second set. However, across the board, the point estimates do hold, particularly for the sample of country-technology pairs that start off with no patent applications whatsoever (i.e., the same sample used in Panel A of Table 2.

Table D1: Migrant inventors and patent applications new technology subclass

Dependent variable: $patents_{c,p,t+10} > 0 \| patents_{c,p,t} = 0$ **(binary)**									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0318		0.0298	0.0178		0.0160	0.0867		0.0862
	(0.004)***		(0.004)***	(0.011)		(0.015)	(0.005)***		(0.008)***
Emigrant inventors		0.0155	0.0054		0.0128	0.0023		0.0975	0.0010
		(0.003)***	(0.002)***		(0.008)	(0.010)		(0.013)***	(0.015)
Total FDI	0.0003	0.0005	0.0002	0.0006	0.0006	0.0006	-0.0009	-0.0026	-0.0010
	(0.000)	(0.000)*	(0.000)	(0.000)*	(0.000)	(0.000)	(0.000)***	(0.001)***	(0.000)*
Total Trade	0.0028	0.0034	0.0023	0.0040	0.0038	0.0038	-0.0016	-0.0073	-0.0017
	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)**	(0.002)***	(0.001)
N	105304	105304	105304	105304	105304	105304	105304	105304	105304
r2	0.26	0.25	0.26	0.26	0.25	0.26	0.24	0.22	0.24
KP F Stat				60.59	37.42	40.91	754.69	99.14	26.79

This table presents results of the estimation of Specification (1). Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair were zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table D2: Migrant inventors and patent applications growth (SPC)

Dependent variable: Patents application growth (SPC), yearly average						
	OLS		2SLS (IV1)		2SLS (IV2)	
	All	$patents_{c,p,t_o} > 0$	All	$patents_{c,p,t_o} > 0$	All	$patents_{c,p,t_o} > 0$
Immigrant inventors	0.0048	0.0053	-0.0005	0.0034	0.0144	0.0155
	(0.001)***	(0.001)***	(0.004)	(0.003)	(0.002)***	(0.001)***
Emigrant inventors	0.0007	0.0009	0.0038	0.0001	0.0003	0.0003
	(0.000)	(0.000)**	(0.003)	(0.002)	(0.004)	(0.003)
Total FDI	0.0000	0.0000	0.0000	0.0001	-0.0002	-0.0002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)*	(0.000)*
Total Trade	0.0006	0.0004	0.0007	0.0007	-0.0003	-0.0003
	(0.000)***	(0.000)***	(0.000)**	(0.000)***	(0.000)	(0.000)
Baseline patent apps, log	0.0066	0.0000	0.0073	0.0000	0.0052	0.0000
	(0.002)***	(.)	(0.002)***	(.)	(0.003)**	(.)
Previous Exports Log-Growth	-0.6119	0.0000	-0.6122	-0.1843	-0.6129	-0.1633
	(0.026)***	(.)	(0.026)***	(0.013)***	(0.027)***	(0.011)***
Zero Exports in t-1	0.0813	-0.0360	0.0805	0.0000	0.0840	0.0000
	(0.005)***	(0.002)***	(0.005)***	(.)	(0.006)***	(.)
N	123690	105304	123690	105304	123690	105304
r2	0.30	0.27	0.30	0.27	0.30	0.26
KP F Stat			44.26	41.07	20.71	26.81

This table presents results of the estimation of Specification (1). Columns 1-2 show OLS estimations, while columns 3-6 show results for 2SLS regressions. Columns 3-4 use as instrument the 30-year lagged stock of migrants; and columns 5-6 use as instrument the predicted stock of migrant inventors based on push and pull factors. The table reports estimations for an alternative measure for decade-long growth in patenting applications: the symmetric percentage change. This measure of growth is defined for country-technology pairs that start off with zero patent applications in the baseline year. Columns 1, 3, and 5 use all the sample, while columns 2, 4, and 6 limit the sample to those observations with initial value 0 in their patent applications count, similarly to Panel A of Table 2. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

E Excluding controls for trade and FDI

Table E1 presents results of our main specification excluding controls for FDI and trade, as some might regard these as "bad controls" since the presence of migrant inventors could also explain an increased flow in trade and FDI. Our main results are robust to this exercise, and the point estimates remain fairly similar.

[Table E1 about here.]

Table E1: Migrant inventors and patent applications take-offs and growth, excl. controls

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0049		0.0049	0.0089		0.0130	0.0086		0.0110
	(0.001)***		(0.001)***	(0.003)***		(0.005)**	(0.002)***		(0.002)***
Emigrant inventors		0.0019	-0.0001		0.0051	-0.0037		0.0063	-0.0032
		(0.001)***	(0.001)		(0.002)***	(0.003)		(0.001)***	(0.003)
N	105304	105304	105304	105304	105304	105304	105304	105304	105304
r2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
KP F Stat				61.15	63.00	39.02	1432.04	504.58	71.31
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0059		0.0038	0.0208		0.0055	0.0678		-0.0615
	(0.001)***		(0.001)***	(0.008)***		(0.011)	(0.065)		(0.313)
Emigrant inventors		0.0056	0.0039		0.0162	0.0126		0.0160	0.0286
		(0.001)***	(0.002)**		(0.007)**	(0.014)		(0.008)*	(0.063)
Baseline patent apps, log	-0.0179	-0.0176	-0.0178	-0.0185	-0.0175	-0.0177	-0.0202	-0.0175	-0.0151
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.003)***	(0.002)***	(0.013)
Previous Exports Growth	-0.1424	-0.1438	-0.1424	-0.1354	-0.1413	-0.1396	-0.1133	-0.1414	-0.1673
	(0.039)***	(0.039)***	(0.039)***	(0.041)***	(0.040)***	(0.037)***	(0.035)***	(0.039)***	(0.127)
Zero Exports in t-1	-0.0016	-0.0014	-0.0014	-0.0015	-0.0007	-0.0009	-0.0010	-0.0008	-0.0007
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
N	18349	18349	18349	18349	18349	18349	18349	18349	18349
r2	0.55	0.55	0.55	0.54	0.54	0.54	0.37	0.54	0.40
KP F Stat				15.90	16.57	7.35	0.35	12.60	0.03

This table presents results of the estimation of Specification (1), excluding controls for trade and FDI. Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair was zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

F Excluding China from the sample

Table F1 presents results of our main specification excluding China, as an attempt to rule out a confounding story that attributes our results to intellectual property stealing facilitated by migrant networks. We choose to exclude China because (1) it represents an important share of migrants due to its size and (2) it is known to have weaker intellectual property protection. Our main results are robust to this exercise.

[Table F1 about here.]

Table F1: Migrant inventors and patent applications take-offs and growth, excl. China

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS					
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0048		0.0048	0.0106		0.0126	0.0097		0.0116
	(0.001)***		(0.001)***	(0.004)**		(0.005)**	(0.002)***		(0.003)***
Emigrant inventors		0.0016	-0.0001		0.0058	-0.0025		0.0094	-0.0038
		(0.001)**	(0.001)		(0.003)**	(0.004)		(0.002)***	(0.004)
Total FDI	-0.0001	-0.0001	-0.0001	-0.0003	-0.0002	-0.0002	-0.0002	-0.0004	-0.0001
	(0.000)	(0.000)	(0.000)	(0.000)**	(0.000)*	(0.000)	(0.000)***	(0.000)***	(0.000)
Total Trade	0.0001	0.0003	0.0002	-0.0003	-0.0002	-0.0002	-0.0003	-0.0007	0.0001
	(0.000)	(0.000)**	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)*	(0.000)**	(0.000)
N	104274	104274	104274	104274	104274	104274	104274	104274	104274
r2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
KP F Stat				60.09	37.10	40.12	732.02	95.33	24.65
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS					
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0048		0.0039	0.0210		0.0084	-0.0140		-0.0019
	(0.001)***		(0.001)***	(0.009)**		(0.010)	(0.038)		(0.046)
Emigrant inventors		0.0040	0.0026		0.0214	0.0166		0.0189	0.0184
		(0.001)***	(0.001)**		(0.016)	(0.021)		(0.024)	(0.032)
Total FDI	0.0003	0.0003	0.0003	0.0002	0.0000	0.0000	0.0005	0.0001	0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Total Trade	0.0021	0.0017	0.0011	-0.0033	-0.0067	-0.0072	0.0083	-0.0055	-0.0046
	(0.001)	(0.001)	(0.001)	(0.003)	(0.008)	(0.008)	(0.014)	(0.011)	(0.029)
Baseline patent apps, log	-0.0170	-0.0167	-0.0169	-0.0175	-0.0165	-0.0169	-0.0163	-0.0165	-0.0165
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Previous Exports Growth	-0.1472	-0.1485	-0.1473	-0.1417	-0.1473	-0.1448	-0.1535	-0.1475	-0.1482
	(0.040)***	(0.040)***	(0.040)***	(0.042)***	(0.040)***	(0.038)***	(0.042)***	(0.040)***	(0.046)***
Zero Exports in t-1	-0.0014	-0.0012	-0.0013	-0.0014	-0.0005	-0.0007	-0.0014	-0.0006	-0.0007
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
N	18077	18077	18077	18077	18077	18077	18077	18077	18077
r2	0.53	0.53	0.53	0.52	0.52	0.52	0.52	0.52	0.52
KP F Stat				8.22	5.74	7.34	0.46	2.41	0.20

This table presents results of the estimation of Specification (1), excluding China from our sample. Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair was zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

G Using alternative patents data

Table G1 replicates the results of Tables 2, but using EPO patent applications as the main source of data, following the description in Section 2.3 (as opposed to USPTO patent applications, used in our main results). Our main results hold with these data.

[Table G1 about here.]

Along the same lines, Table G2 replicates the results using patent applications based on PCT records as the main source of data, following the description in Section 2.3. Our main results hold with these data.

[Table G2 about here.]

In addition, Table G3 replicates our main results using granted patents–as opposed to patent applications–to construct our dependent variables, as well as those control variables that are based on patents data. In the table, we document that our main results are robust to using granted patents as opposed to patent applications.

We consider these results of high relevance as, even when there could be a significant time gap between the time of patent application and the moment it is granted, the take-offs and accelerations we document in innovation are also present when focusing on patents that have gone through the process imposed by patent agencies that certify the innovation.

[Table G3 about here.]

Table G1: Migrant inventors and patent applications take-offs and growth (EPO)

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0040		0.0038	0.0043		0.0008	0.0067		0.0036
	(0.001)***		(0.001)***	(0.002)*		(0.003)	(0.001)***		(0.002)*
Emigrant inventors		0.0019	0.0007		0.0052	0.0046		0.0104	0.0060
		(0.001)***	(0.001)		(0.002)**	(0.003)*		(0.002)***	(0.003)**
Total FDI	0.0001	0.0002	0.0001	0.0001	0.0000	0.0000	0.0001	-0.0002	-0.0001
	(0.000)**	(0.000)**	(0.000)*	(0.000)*	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total Trade	0.0001	0.0002	0.0000	0.0001	-0.0002	-0.0002	-0.0001	-0.0009	-0.0007
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)***	(0.000)**
N	105039	105039	105039	105039	105039	105039	105039	105039	105039
r2	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
KP F Stat				59.94	38.38	40.35	938.44	93.47	26.07
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0059		0.0048	0.0128		0.0047	-0.0139		-0.0377
	(0.001)***		(0.001)***	(0.007)*		(0.006)	(0.018)		(0.212)
Emigrant inventors		0.0057	0.0042		0.0147	0.0118		0.0259	-0.0509
		(0.002)***	(0.001)***		(0.007)**	(0.007)		(0.028)	(0.460)
Total FDI	-0.0002	-0.0002	-0.0002	-0.0002	-0.0003	-0.0003	-0.0002	-0.0005	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.005)
Total Trade	0.0073	0.0059	0.0044	0.0034	-0.0018	-0.0020	0.0187	-0.0113	0.0756
	(0.003)***	(0.003)**	(0.003)*	(0.003)	(0.005)	(0.005)	(0.013)	(0.023)	(0.506)
Baseline patent apps, log	-0.0209	-0.0207	-0.0208	-0.0211	-0.0205	-0.0207	-0.0205	-0.0203	-0.0208
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.001)***	(0.002)***	(0.005)***
Previous Exports Growth	-0.1239	-0.1256	-0.1242	-0.1219	-0.1254	-0.1241	-0.1299	-0.1252	-0.1380
	(0.025)***	(0.026)***	(0.025)***	(0.026)***	(0.026)***	(0.026)***	(0.027)***	(0.026)***	(0.079)*
Zero Exports in t-1	-0.0006	-0.0004	-0.0004	-0.0006	0.0000	-0.0001	-0.0008	0.0006	-0.0035
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.024)
N	18623	18623	18623	18623	18623	18623	18623	18623	18623
r2	0.47	0.47	0.47	0.47	0.47	0.47	0.45	0.46	0.24
KP F Stat				7.74	10.57	9.91	1.49	2.40	0.01

This table presents results of the estimation of Specification (1), using patent applications reported by the EPO to construct the dependent variable as well as patent-relevant control variables. Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair was zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

Table G2: Migrant inventors and patent applications take-offs and growth (PCT)

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0058		0.0047	0.0101		0.0075	0.0095		0.0072
	(0.001)***		(0.001)***	(0.003)***		(0.004)*	(0.002)***		(0.003)**
Emigrant inventors		0.0044	0.0027		0.0078	0.0030		0.0131	0.0046
		(0.001)***	(0.001)***		(0.003)***	(0.003)		(0.003)***	(0.004)
Total FDI	0.0001	0.0001	0.0000	-0.0000	-0.0000	-0.0001	-0.0000	-0.0003	-0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)**	(0.000)
Total Trade	0.0002	0.0001	-0.0000	-0.0001	-0.0003	-0.0002	-0.0000	-0.0008	-0.0004
	(0.000)**	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)***	(0.000)
N	106327	106327	106327	106327	106327	106327	106327	106327	106327
r2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
KP F Stat				54.56	41.29	35.09	664.82	100.84	27.92
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0057		0.0044	0.0145		0.0040	-0.0256		-0.0239
	(0.002)***		(0.002)**	(0.006)**		(0.005)	(0.023)		(0.022)
Emigrant inventors		0.0060	0.0046		0.0171	0.0150		0.0160	0.0068
		(0.001)***	(0.001)***		(0.006)***	(0.008)*		(0.019)	(0.021)
Total FDI	0.0003	0.0002	0.0002	0.0002	0.0000	0.0000	0.0005	0.0001	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Total Trade	0.0029	0.0016	0.0006	-0.0010	-0.0053	-0.0057	0.0167	-0.0046	0.0117
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.011)	(0.011)	(0.016)
Baseline patent apps, log	-0.0248	-0.0247	-0.0247	-0.0248	-0.0245	-0.0245	-0.0247	-0.0245	-0.0246
	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.002)***	(0.003)***	(0.002)***	(0.003)***
Previous Exports Growth	-0.0684	-0.0687	-0.0692	-0.0700	-0.0711	-0.0714	-0.0629	-0.0709	-0.0647
	(0.023)***	(0.023)***	(0.023)***	(0.023)***	(0.022)***	(0.022)***	(0.026)**	(0.022)***	(0.024)***
Zero Exports in t-1	-0.0052	-0.0052	-0.0051	-0.0047	-0.0048	-0.0047	-0.0069	-0.0048	-0.0066
	(0.002)**	(0.002)**	(0.002)**	(0.002)**	(0.002)*	(0.002)*	(0.002)***	(0.003)*	(0.002)***
N	17323	17323	17323	17323	17323	17323	17323	17323	17323
r2	0.59	0.59	0.60	0.59	0.59	0.59	0.57	0.59	0.57
KP F Stat				7.98	17.58	8.61	1.49	0.84	0.31

This table presents results of the estimation of Specification (1), using patent applications reported by the PCT to construct the dependent variable as well as patent-relevant control variables. Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair was zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis. $^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

Table G3: Migrant inventors and granted patent take-offs and growth (USPTO)

Panel A - Dependent variable: Patent applications take-off (binary)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0048		0.0045	0.0084		0.0090	0.0113		0.0106
	(0.001)***		(0.001)***	(0.004)**		(0.005)*	(0.002)***		(0.002)***
Emigrant inventors		0.0025	0.0009		0.0051	-0.0008		0.0135	0.0016
		(0.001)***	(0.001)		(0.002)**	(0.003)		(0.003)***	(0.004)
Total FDI	-0.0000	0.0000	-0.0000	-0.0001	-0.0001	-0.0001	-0.0002	-0.0004	-0.0002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)*	(0.000)***	(0.000)
Total Trade	0.0001	0.0002	0.0000	-0.0002	-0.0001	-0.0001	-0.0004	-0.0012	-0.0005
	(0.000)	(0.000)*	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)***	(0.000)***	(0.000)
N	106620	106620	106620	106620	106620	106620	106620	106620	106620
r2	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.08	0.09
KP F Stat				60.25	37.25	41.17	738.22	95.44	26.70
Panel B - Dependent variable: Patent applications growth (CAGR)									
	OLS			2SLS (IV1)			2SLS (IV2)		
	est1	est2	est3	est4	est5	est6	est7	est8	est9
Immigrant inventors	0.0026		0.0012	0.0167		-0.0012	0.1321		0.1240
	(0.001)**		(0.001)	(0.010)*		(0.009)	(0.163)		(0.162)
Emigrant inventors		0.0038	0.0033		0.0198	0.0206		0.0275	0.0184
		(0.001)***	(0.002)**		(0.014)	(0.019)		(0.034)	(0.070)
Total FDI	0.0004	0.0004	0.0004	0.0003	0.0001	0.0001	-0.0003	0.0000	-0.0005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
Total Trade	0.0002	-0.0005	-0.0007	-0.0037	-0.0066	-0.0065	-0.0357	-0.0095	-0.0404
	(0.001)	(0.001)	(0.001)	(0.002)	(0.005)	(0.005)	(0.046)	(0.011)	(0.048)
Baseline patent apps, log	-0.0170	-0.0168	-0.0168	-0.0175	-0.0164	-0.0164	-0.0215	-0.0163	-0.0208
	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.002)***	(0.006)***	(0.003)***	(0.007)***
Previous Exports Growth	-0.1291	-0.1300	-0.1298	-0.1275	-0.1324	-0.1327	-0.1144	-0.1336	-0.1181
	(0.044)***	(0.044)***	(0.044)***	(0.045)***	(0.041)***	(0.039)***	(0.048)**	(0.043)***	(0.053)**
Zero Exports in t-1	-0.0025	-0.0022	-0.0023	-0.0027	-0.0013	-0.0013	-0.0046	-0.0009	-0.0035
	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.008)
N	17032	17032	17032	17032	17032	17032	17032	17032	17032
r2	0.50	0.50	0.50	0.49	0.49	0.49	-0.30	0.47	-0.30
KP F Stat				10.74	9.75	7.46	0.35	0.50	0.35

This table presents results of the estimation of Specification (1), using granted patents by the USPTO to construct the dependent variable as well as patent-relevant control variables. Columns 1-3 show OLS estimations, while columns 4-9 show results for 2SLS regressions. Columns 4-6 use as instrument the 30-year lagged stock of migrants; and columns 7-9 use as instrument the predicted stock of migrant inventors based on push and pull factors. Panel A presents results for take-offs in patenting applications (limiting the sample to cases where the initial patent applications for that country-technology pair was zero), while Panel B estimates future CAGR in patent applications for country-technology pairs that already had some patenting recorded in the baseline year. All specifications include country-year and technology-year fixed effects. SEs clustered at the country level are presented in parenthesis.

$^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$