2015

International knowledge flows and technological advance: the role of migration

Kacey N. Douglas
University of Nebraska-Lincoln, kacey.douglas@gmail.com

Douglas *IZA Journal of Migration* (2015) 4:13
DOI 10.1186/s40176-015-0037-8

IZA Journal of Migration
a SpringerOpen Journal

ORIGINAL ARTICLE Open Access

International knowledge flows and technological advance: the role of migration

Kacey N Douglas

Correspondence:
kacey.douglas@gmail.com
University of Nebraska-Lincoln, P.O. Box 880489, Lincoln, NE 68588-0489, USA

Abstract

This paper investigates whether knowledge flows from host to source country as a result of migration, alleviating the negative effects associated with outward migration. Using a fixed effects Poisson regression, patent citations are used as a proxy for knowledge flows and regressed on immigration and other control variables; the effect of immigration on patent citations is found to be positive and statistically significant. Additionally, the coefficient on immigration is found to be robust to different parameter changes in the model. These results suggest that reverse knowledge flows from outward migration help mitigate negative effects of outward migration on source countries.

JEL Classification Numbers: F22; O30; O15

Keywords: Immigration; Patent citations; Knowledge flows; Brain drain

Introduction

Standard growth theory points out that technological progress is critical for achieving sustainable economic growth. However, the existing immigration literature analyzing the costs and benefits of immigration has often ignored the possibility that immigration affects the growth of technology in both source and destination countries. Perhaps this failure is due to the fact that the traditional labor market models on which most immigration analysis is based do not address technological change. This paper looks outside the realm of traditional labor market models of immigration to investigate whether immigration can create technological advance and thus long-run economic growth.

Traditional labor market models of immigration conclude that international migration leads to an increase in income in the host country and a decrease in the source country. However, if this migration could create knowledge flows from host to source countries, the detrimental effects on source countries associated with outward migration may be less than expected.

There are several avenues through which migration can send knowledge from host countries back to source countries. Mayr and Peri (2008) suggest that technology flows back to source countries when migrants return; for example, they provide evidence that highly skilled migrants increasingly migrate temporarily and, therefore, bring back with them the knowledge they acquire abroad. However, with the rapid expansion in information and communications technology (ICT), migrants no longer need to return home in order to influence technologies in their homelands. Migrants in destination countries create "diaspora networks" with the purpose of sharing knowledge with their source countries. In 2005, UNESCO's Diaspora Knowledge Network (DKN) project was initiated in order to strengthen these networks and their abilities to utilize ICT (Grossman 2010). Though some groups have been more successful than others, the diaspora networks have the ability to advance social and economic development in the home countries (de Haas 2006). Saxenian (2002) also argues that immigration can benefit the source country because immigrants in host countries support their counterparts at home. Using the case of Silicon Valley, Saxenian shows that the numerous ethnic groups, who account for a large number of the Valley's highly skilled workers, maintain relationships, both social and professional, with their professional colleagues at home, creating information flows back to the source country.

This paper extends the existing literature to investigate empirically the relationship between international migration flows and international knowledge flows. Specifically, patent citations are used to measure the relationship between knowledge flows from the United States to a sample of foreign countries and US immigration. First, the methodology for capturing knowledge flows is explained. Next, the econometric model and the data are described. Then, the results of the regression analysis are presented and discussed. Next, sensitivity analyses are performed on the model to determine robustness. Finally, all results are discussed and future refinements and extensions to the hypothesis, the model and the paper are explored.

Brain drain migration

This paper investigates the hypothesis that migration creates knowledge flows from host to source countries. If true, such knowledge transfers and the potential positive growth effects would mitigate some of the detrimental effects of out migration on source countries. According to the literature, this detrimental effect is exacerbated when the migration is "brain drain" migration, namely, highly skilled labor leaving one country to find better economic opportunities in another. As a result, increased knowledge flows offsetting brain drain migration specifically would have a greater impact on reducing the welfare losses associated with outward migration. Thus, we may ask whether the sample countries in this paper have experienced outward brain drain migration.

As a partial answer to this question, Saint-Paul (2004) uses US and European census data to reveal that the brain drain process is indeed occurring from Western Europe to the US; Europeans living in the US are vastly outperforming both their American and European counterparts. To this point, the table below shows the percentage of the expatriate population with tertiary education versus the corresponding percentage in home country and the whole US in 1990 and 2000.

According to Table 1, Europeans living in the US are more likely than their US counterparts and approximately twice as likely as their European counterparts to have tertiary educational attainment. Furthermore, the table below shows the percentage of European expatriates with a Ph.D. as compared to the percentage of the whole US population in 1990 and 2000 (Table 2).

Table 1 Percentage of population with tertiary education (Saint-Paul 2004)

Country	1990		2000	
	In United States	In Home Country	In United States	In Home Country
Belgium	47.6 %	17 %	59.6 %	26 %
France	42.7	14	56.1	24
Germany	34.6	17	41.9	28
Great Britain	38.9	15	49.5	25
Italy	17.1	6	25.7	13
Spain	30.6	9	44.1	21
United States	29.7	N/A	33.8	N/A

Table 2 Percentage of Europeans in US with a Ph.D. (Saint-Paul 2004)

Country	1990	2000
Belgium	4.33 %	5.78 %
France	3.1	4.9
Germany	1.72	2.39
Great Britain	3.2	3.9
Italy	0.96	2.0
Spain	2.7	4.6
United States	0.82	0.98

The above table reveals that European expatriates are more likely to hold Ph.D.s than the US as a whole.

Likewise, Murakami (2010) suggests that Japan is also suffering from brain drain migration to the US. He writes that "…a considerable number of Japanese researchers and engineers are moving overseas, primarily to the United States." He goes on to add that "…the number of Japanese individuals living in the United States who have an undergraduate or higher level of education, and who have a degree in a field related to science or engineering is as high as 59,400."[1] Thus, the results of this analysis are very relevant in that migration leading to knowledge transfers can, in fact, help mitigate the detrimental effects associated with the brain drain process.

Patent citations

To investigate empirically the relationship between migration and knowledge flows, one must first capture some measurement of the flow of knowledge. Knowledge flows, however, are difficult to measure; they rarely leave a paper trail to follow. Different methods have been employed in an attempt attempt to quantify these flows. Regets (2001) uses the existence of international coauthors to measure, in some sense, international knowledge flows. Regets finds that the percentage of a foreign country's internationally coauthored articles with the US increases with the amount of US doctorates received by immigrants from that country. His findings suggest that migration increases knowledge flows. Coe and Helpman (1995) use R&D expenditures as a proxy for the stock of knowledge with the intent to measure its effects on productivity. They find that foreign R&D increases domestic productivity, indicating some international flow of knowledge.

Another method, which is employed in this paper, uses patents; patents contain information that can be used to measure knowledge flows. A patent creates a temporary property right over some piece of knowledge, technique, process, or method; it is granted to an inventor or applicant by a sovereign state, in most cases a country. Often, a patent is an extension of previously patented technology. If so, that subsequent patent (the *citing patent*) must cite the previous patent (the *originating patent*) upon which it builds—just as an author of an academic article must cite previous knowledge used. Each patent is recorded as a public document containing detailed information regarding the inventor, including their geographic location. By examining the location of the inventor of both the originating patent and the citing patent, it becomes possible to ascertain the path of knowledge flows—from the location of the originating-patent inventor to the location of the citing-patent inventor.

Thus, patent citations can be used as a proxy for knowledge flows. Jaffe et al. (1993) use patent citation data to measure technology flows within North America. Their study finds that citing patents are more likely to occur in the same geographic location as the originating patent, indicating that knowledge flows are geographically localized. Agrawal et al. (2003) use the same methodology as Jaffe et al. to capture knowledge flows in the US and Canada, examining patent activity in areas where inventors previously resided. They find that patents are cited disproportionately where the inventor receiving the patent previously resided, revealing that knowledge flows do result from migration.

Though citing patents can be used to measure knowledge flows within a country or region, they have been more difficult to trace internationally. However, a global Patent Cooperation Treaty (PCT), concluded in Washington Diplomatic Conference on the Patent Cooperation Treaty 1970, made great strides in providing transparency of patents' information. The PCT currently has 148 contracting member states, and any patent application from a member state is required to include citations of all previous patents upon which this new patent builds. Though patent protection can only be offered within a country, the PCT requires the search for all previous patents be performed on an international scale.[2] As a result, an inventor seeking to patent in a member country must cite all prior patents, including those from other member countries.

The widespread membership of the PCT has successfully ensured the existence of international patent citations. However, while this documentation of international citations exists, it is not widely available. Many patent databases contain only citations of national patents. Even those that do contain international information often have only a limited number of countries included in their database. These data limitations quickly become cumbersome for empirical analysis. As a result, little literature exists investigating the correlation between international migration and international knowledge flows using patent citations as a proxy for knowledge flows.

Fortunately, the European Patent Office (EPO) operates a database and patent search engine entitled Espacenet, which contains over 70 million patent documents from 1836 to the present. Each patent document on Espacenet contains information on "citing documents", which include any citing patents. The citing patents contain information on the country of residence of the associated inventor. If a US patent is cited by a foreign inventor, this is considered a forward citation. This paper uses this database to obtain forward citations as a proxy for knowledge flows from the inventor of the originating patent to the inventor of the citing patent.

Modeling reverse-knowledge flows

In order to test the importance of immigration as a determinant of knowledge flows, we specify an empirical model as follows:

Citation	Number of times a US patent is cited by a unique patent with an inventor from country *j* at time *t*

There are undoubtedly many factors that affect the dependent variable, some specific to the country involved and others which are more general. For purposes of this study, I consider seven.[3] The independent variables considered here are:

GDP	Gross domestic product in current US dollars in country *j* at time *t*
Trade	Imports of US goods plus exports to US in millions of current US dollars in country *j* at time *t*
Patent Stock Ag	Sum of total *agricultural* patents and patent applications in country *j* at time *t*
Education	Percentage of the student aged population enrolled in tertiary education in country *j* at time *t*
Inward FDI	Inward foreign direct investment stock in millions of current US dollars from country *j* at time *t* in US
Outward FDI	Outward foreign direct investment stock in millions of current US dollars from US in country *j* at time *t*

And finally, the independent variable reflecting my hypothesis:

Immigration	Sum of total *employment-based* immigration (in thousands) to US from country *j* for five years prior to time *t*

For variable data sources, see Appendix (Table 3).

A positive coefficient for an independent variable suggests that an increase in the value of that variable increases the amount of forward citations, ceteris paribus. The sign of GDP is expected to be positive; prior literature has shown that, holding other variables constant, patents issued are positively correlated with GDP (Florida 2005). A country with a higher GDP issues more patents, cites more patents in general, and most likely cites more US patents as well. A positive correlation between trade

Table 3 Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Immigration	237540	7.951135	10.32182	0.614	48.447
GDP	267072	1.13e + 12	1.28e + 12	1.22e + 11	5.46e + 12
Trade	267072	43714.95	51848.34	1330.4	211403.8
Patent Stock Ag	267072	1152.221	2135.329	9	9761
Education	236256	58.4958	15.43909	20.09968	95.01728
Inward FDI	263220	79178.04	94175.33	−41	447529
Outward FDI	265788	71754.97	103683.6	533	514689

(TRADE) and patent citations has been shown empirically by numerous authors, including Sjöholm (1996), Hu and Jaffe (2003) and MacGarvie (2005).

The argument of a positive correlation between patent stock and citations is analogous to that of GDP's correlation with citations; more patent activity leads to more patent citations in general, including citations of US patents. The stock of agricultural patents and patent applications (PATENT STOCK AG) is included in regression because the originating patents in the sample are agricultural patents; a larger stock of patents may lead to increased patent citations. The positive correlation between education (EDUCATION) and patent citations is expected because higher education leads to a more skilled labor force, which would then be more likely to create technological advance via patents. Both measures of Foreign Direct Investment (FDI) are expected to be positively correlated with foreign citation of US patents, because FDI has been shown to be an avenue through which knowledge flows from source to host country and vice versa (Saggi 2002, Hu and Jaffe 2003, MacGarvie 2005).

The hypothesis explored here is that immigration is also expected to have a positive sign for the reasons outlined above, in particular the host-to-source-country back-linkages resulting from immigration. Specifically, the sign on immigration (IMMIGRATION) is expected to be positive.

Methodology

This paper aims to measure the flow of knowledge from the US outward. To begin, a sample of US patents is chosen and all citing patents from foreign inventors are found.[4] To do so, each US patent must be searched individually on Espacenet. Then, each citing patent must be investigated in order to gather the necessary information needed for the dataset—namely, country of residence of the inventor and date. Because of the time commitment required to gather this information, it is necessary to narrow down the beginning sample of US patents to a manageable number. Patenting activity in the US is immense; in 1998 alone 163,204 patents were granted. Therefore, a class, or subset of patents, needed to be chosen.

One might first consider classes of patents where patent activity is most intensive, such as high-tech patents. Another method would be to consider patents in fields where immigration has had a strong impact. Studies have shown that migrants have had a particularly strong impact in Science, Technology, Engineering, and Mathematics (STEM) fields and occupations (Kerr 2013). The methodology for choosing the class of patents in this paper is as follows: reverse knowledge flows would tend to be most effective in areas where a large portion of economic activity presides. Developing countries rely on the agricultural sector as both an important source of viability and income. As a result, advances in technology related to agriculture would be highly beneficial for developing countries. Though the majority of the foreign countries in this study are developed countries due to data restrictions, evidence of reverse knowledge flows in the agricultural sector will be of particular import for developing countries. Because the out-migration of highly educated people may be most detrimental to developing countries, the results of this paper could have an even greater impact for lesser developed countries suffering from brain drain migration. So, the knowledge flows related to agricultural are of most interest.

Thus, a sample was chosen of US patents consisting of 1,284 patents from US Patent Class entitled "Plant Husbandry" granted to US inventors from 1998 to 2002.[5] These

1,284 patents represent roughly 2.5% of all US patents granted to US investors over this time period. Plant Husbandry is defined by the US Patent and Trademark Office (USPTO) as " … the parent class for apparatus and processes employed in treating the earth and its products and includes all inventions relating thereto that have not been especially provided for in other classes." This classification contains agricultural patents.

The dependent variable in the following regression analysis is number of times a US patent is cited by a unique patent with an inventor from a given country in a given year. Of the 1,284 US patents, 473 (or 37%) have forward citations. The share of forward citations by inventor country is listed below (Table 4):

The independent variables consist of factors that could affect the frequency with which these citations occur. That is, the independent variables explain how often a US patent is cited by a patent with an inventor from a given sample country in a given year. The above table also includes the share of employment-based migration to the US by inventor country. Graphically, this data is illustrated below (Fig. 1):

The simple scatter plot above shows a clear positive correlation between migration and forward citations, as predicted in the hypothesis of this paper. An empirical model is developed in the next section to further investigate this relationship.

Model

Because the dependent variable in the regression is a nonnegative count variable with no theoretical upper bound—it takes on integer values greater than or equal to zero—the most appropriate econometric model for the analysis conducted in this paper is the fixed effects (FE) Poisson regression model developed by Hausman, Hall, and Griliches (Hausman et al. 1984). Consider the following linear model for T time periods:

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + c_i + u_{it}, \quad t = 1, 2, \ldots, T; i = 1, 2, \ldots, N, \tag{1}$$

Table 4 Share of citations, migration by country

Country	Share of citations	Share of migration
Australia	8.68 %	4.98 %
Belgium	3.02	1.44
Switzerland	1.89	2.02
Germany	20	16.53
Denmark	1.76	1.32
Spain	7.17	3.40
Finland	1.38	0.96
France	15.09	7.57
Great Britain	18.62	32.95
Greece	0.50	2.74
Japan	10.57	15.50
Netherlands	10.94	3.04
Turkey	0.38	7.56

Notes: Column 2 reports forward citations of US by country of inventor as a percentage of total forward citations. Column 3 reports employment-based immigration from 1990 to 2010 by country as a percentage of total employment-based immigration from all of the sample countries from 1990 to 2010

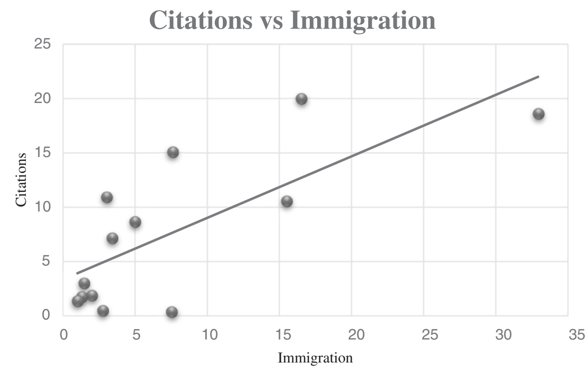

Fig. 1 *Notes:* The above is a scatter plot of information contained in Table 4. Citations reports forward citations of US by country of inventor as a percentage of total forward citations. Immigration reports employment-based immigration from 1990 to 2010 by country as a percentage of total employment-based immigration from all of the sample countries from 1990 to 2010

where c_i is an unobserved, time invariant effect associated with each US patent. In addition, $\mathbf{x}_{it}$ is the vector of independent variables associated with patent i at time t. If y given under $\mathbf{x}$ is distributed as Poisson, the density function is given as:

$$\begin{aligned} f(y_{it}|\mathbf{x}_{it},\ c_i) &= \mathrm{P}(\mathrm{Y}_{it} =\ y_{it}|\mathbf{x}_{it},\ c_i) =\ \exp\left\{-\exp(a_i + \mathbf{x}_{it}\boldsymbol{\beta})\right\} \exp(a_i + \mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}}/y_{it}! \\ &= \frac{1}{y_{it}!} \exp\{-\exp(a_i)\exp(\mathbf{x}_{it}\boldsymbol{\beta}) + a_i y_{it}\} \exp(\mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}} \end{aligned} \tag{2}$$

If $\mathrm{E}(y_t|\mathbf{x}_1,\ \ldots,\ \mathbf{x}_T,\ c) = \mathrm{E}(y_t|\mathbf{x}_t,\ c)$ under the assumption of exogeneity, the joint probability density function within a panel can be written as:

$$\begin{aligned} f(y_i|\mathbf{X}_i,\ c_i) &= \mathrm{P}(\mathrm{Y}_{i1} = y_{i1},\ \ldots,\ \mathrm{Y}_{iT} = y_{iT}|\mathbf{X}_i,\ c_i) \\ &= \prod\nolimits_{t=1}^{T} \frac{1}{y_{it}!} \exp\{-\exp(a_i)\exp(\mathbf{x}_{it}\boldsymbol{\beta}) + a_i y_{it}\} \exp(\mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}} \\ &= \left(\prod\nolimits_{t=1}^{T} \frac{\exp(\mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}}}{y_{it}!}\right) \exp\left\{-\exp(a_i)\sum\nolimits_t \exp(\mathbf{x}_{it}\boldsymbol{\beta}) + a_i \sum\nolimits_t y_{it}\right\} \end{aligned} \tag{3}$$

The conditional likelihood function is obtained using a joint probability distribution conditional on the sum of outcomes across t:

$$\begin{aligned} &\mathrm{P}\Big(\mathrm{Y}_{i1} = y_{i1},\ \ldots,\ \mathrm{Y}_{it} = y_{it}|\mathbf{X}_i, c_i,\ \Big(\sum\nolimits_t \mathrm{Y}_{it} = \sum\nolimits_t y_{it}|\mathbf{X}_{i,}c_i\Big)\Big) \\ &= \frac{\left[\left(\prod\nolimits_{t=1}^{T} \frac{\exp(\mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}}}{y_{it}!}\right) \exp\left\{-\exp(a)\sum\nolimits_t \exp(\mathbf{x}_{it}\boldsymbol{\beta}) + a_i \sum\nolimits_t y_{it}\right\}\right]}{\frac{1}{\left(\sum\nolimits_t y_{it}\right)!} \exp\left\{-\exp(a_i)\sum\nolimits_t \exp(\mathbf{x}_{it}\boldsymbol{\beta}) + a_i \sum\nolimits_t y_{it}\right\}\left\{\sum\nolimits_t \exp(\mathbf{x}_{it}\boldsymbol{\beta})\right\}^{\sum\nolimits_t y_{it}}} \\ &= \left(\sum\nolimits_t y_{it}\right)! \prod\nolimits_{t=1}^{T} \frac{\exp(\mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}}}{y_{it}!\left\{\sum\nolimits_{r=1}^{T} \mathbf{x}_{ir}\boldsymbol{\beta}\right\}^{y_{it}}} \end{aligned} \tag{4}$$

The FE Poisson estimator, $\hat{\boldsymbol{\beta}}_{FEP}$, is defined as the estimator that maximizes the conditional log likelihood function:

$$\begin{aligned} l = & \log\prod\nolimits_{i=1}^{N}\left[\left(\sum\nolimits_{t}y_{it}\right)!\prod\nolimits_{t=1}^{T}\frac{\exp(\mathbf{x}_{it}\boldsymbol{\beta})^{y_{it}}}{y_{it}!\left\{\sum\nolimits_{r}\exp(\mathbf{x}_{ir}\boldsymbol{\beta})\right\}^{y_{it}}}\right] \\ = & \log\prod\nolimits_{i=1}^{N}\left\{\frac{\left(\sum\nolimits_{t}y_{it}\right)!}{\coprod\nolimits_{t=1}^{T}y_{it}!}\prod\nolimits_{t=1}^{T}p_{it}^{y_{it}}\right\} \\ = & \sum\nolimits_{i=1}^{N}\left\{\log\Gamma\left(\sum\nolimits_{t=1}^{T}y_{it}+1\right)-\sum\nolimits_{t=1}^{T}\log\Gamma(y_{it}+1)+\sum\nolimits_{t=1}^{T}y_{it}\log p_{it}\right\}, \end{aligned} \tag{5}$$

where $p_{it} = {}^{\exp^{\mathbf{x}_{it}\beta}}\Big/_{\sum_r \exp^{\mathbf{x}_{ir}\boldsymbol{\beta}}}$, and y given under $\mathbf{x}$ is distributed as Poisson.

That is, $\hat{\boldsymbol{\beta}}_{FEP}$ will be chosen to solve the following equation:

$$\sum\nolimits_{i=1}^{N}\left(\partial l_i\left(\hat{\boldsymbol{\beta}}_{FEP}\right)\Big/_{\partial\hat{\boldsymbol{\beta}}_{FEP}}\right) = 0, \tag{6}$$

where:

$$\partial l_i\left(\hat{\boldsymbol{\beta}}_{FEP}\right)\Big/_{\partial\hat{\boldsymbol{\beta}}_{FEP}} = \sum\nolimits_{t=1}^{T}y_{it}\left[\left(\frac{\partial p_{it}}{\partial\beta_{FEP}}\right)'\Big/_{p_{it}}\right] \tag{7}$$

This estimation method has the attractive robustness property that, under only the assumption of exogeneity, the fixed effects Poisson (FEP) estimator, $\hat{\boldsymbol{\beta}}_{FEP}$, is consistent.

Data

The following model was used for the regression:

$$\begin{aligned} citation_{ijt} = & \, immi\ddot{g}ration_{ijt}\beta_1 + G\ddot{D}P_{ijt}\beta_2 + tr\ddot{a}de_{ijt}\beta_3 + patent\ \ddot{s}tock\ ag_{ijt}\beta_4 \\ & + educ\ddot{a}tion_{ijt}\beta_5 + inwar\ddot{d}\ FDI_{ijt}\beta_6 + outwar\ddot{d}\ FDI_{ijt}\beta_7 + u_{ijt} \end{aligned} \tag{8}$$

Where $i = 1, 2, \ldots, 1284$ *US patents*

$j = 1, 2, \ldots, 13$ *Countries*

and $t = 1995, 1996, \ldots, 2010$ *Years*[6]

Due to data restrictions, only 13 countries were included in the set; only information from these countries' patent offices are contained in Espacenet (for simplicity's sake, these countries will be called member countries). For example, the US is a member country. This means that information from the USPTO is included in the search engine. Thus, Espacenet will have documentation of forward citations for US patents, including inventors worldwide who have sought patent protection in the US, in any of the other member countries, or in any of the international patent application organizations included in the Espacenet database. China is not a country whose patent office's information is included in Espacenet. Thus, no citing patents from China will be revealed through an Espacenet search. In other words, citing patents from Chinese inventors will only be found on Espacenet if these inventors are applying for patents in one of the member countries or international patent application organizations. It is highly likely that a nontrivial number of Chinese inventors seek patent protection in China. Thus, a large amount of citing patents from Chinese inventors will not be revealed via an Espacenet patent search. So, it would not be wholly representative to include only forward citations from Chinese inventors seeking patent protection in member countries or the aforementioned international patent application organizations. Therefore, only citing patents from inventors residing in sample countries are included. Results of the above regression are discussed in the following section.

Results

Table 5 summarizes the results obtained from performing conditional MLE on the FE Poisson regression model.[7] The coefficients can be interpreted in the following manner: for a one unit change in the independent variable, that variable's coefficient is equal to the percentage change in the predicted amount of forward citations, holding all other independent variables constant. As predicted, education and inward FDI are positive and statistically significant. Additionally, the coefficient on the variable of interest, immigration, is positive and statistically significant. The coefficient on immigration shows that an increase of 1,000 immigrants into the US from country j in the five years prior to year t is associated with a 0.03% increase in the number of forward citations with inventors in country j in year t. To put this into perspective, approximately 41,000 people emigrated from Great Britain to the US between 2001 and 2005. According to the coefficient on immigration, this would result in a 1.23% increase in the number of forward citations with inventors from Great Britain in 2006. Thus, immigration does have a positive and statistically significant effect on the amount of forward

Table 5 FE poisson regression

Variable	Coefficient (Standard Error)	z-statistic	95 % Confidence Interval	
GDP	1.12e-14 (1.02e-13)	0.11	−1.88e-13	2.11e-13
Trade	8.58e-07 (3.14e-06)	0.27	−5.29e-06	7.00e-06
Patent Stock Ag	0.0000616 0(.0000455)	1.35	−0.0000276	0.0001509
Education	0.0074407 (0.0017043)	4.37***	0.0041004	0.0107811
Inward FDI	1.50e-06 (6.33e-07)	2.37**	2.58e-07	2.74e-06
Outward FDI	−8.43e-07 (6.15e-07)	−1.37	−2.05e-06	3.62e-07
Immigration	0.0311889 (0.00519)	6.01***	0.0210168	0.0413611

Notes: Observations = 64000. Standard errors are robust to heteroskedasticity
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

citations with foreign inventors; specifically, knowledge flows are positively correlated with return migration flows.

Robustness

Many of the empirical studies only use very specific models with a relatively small number of explanatory variables in order to report a statistically significant relationship between two variables of interest. As a result, the majority of conclusions drawn in the literature are fragile; they depend on the conditioning set of information in the regression model. Thus, two sensitivity analyses are performed below in hopes of providing "full disclosure" and robust results.

First, Leamer and Leonard (1983) argue that the advance of econometric technology has allowed economists to draw conflicting inferences from the same data. They encourage researchers to summarize the entire range of inferences implied by a whole family of alternative models using given data. In effect, they propose is a sensitivity analysis that consists of systematically changing the parameterization of the model and reporting the results. They conduct this analysis by imposing various combinations of exclusion restrictions around one variable of interest and observe whether the coefficient on the variable of interest remains statistically significant and of the same sign. This analysis allows a reporting of results that is much more informative than the results often reported in the literature.

Levine and Renelt (1992) perform a similar analysis to that proposed by Leamer and Leonard. These authors use data regarding the long-run growth rates and a variety of regressors linked to growth in the literature. They run numerous regressions, always including a chosen set of independent variables and alternating a separate set of independent variables for each regression. They find almost all variables of interest fragile, meaning they do not remain the same sign and statistically significant over the range of regressions.

To test whether immigration is fragile, the regression is again run with immigration as the constant independent variable; the remaining independent variables are included interchangeably in sets of three. There are a total of $\binom{6}{3} = 20$ regression models. The results are reported below (Table 6).[8]

The immigration variable in this sensitivity analysis is definitively robust; every regression, regardless of the regressors included, yields a positive and statistically significant immigration correlation coefficient. Thus, one can conclude that there is a robustly positive correlation between citations and immigration, or between knowledge flows and migration flows.

In an additional attempt to increase robustness, one may consider other definitions of immigration. To test whether the definition of immigration affects the results, the regression is again run using the initial covariates and different measures of immigration. The descriptive statistics and regression results are listed below (Tables 7 and 8).

Table 8 reveals that immigration, regardless of how it is defined, has a positive and statistically significant impact on knowledge flows. Thus, immigration is not sensitive to the measurement technique used. This table also reveals that total

Table 6 Sensitivity analysis, regressors

#	Variables In Regression	Coefficient on Immigration (Standard Error)	z-statistic	Obs	Sign	Significant
0	None	0.0315809 (0.0022325)	14.15***	86580	+	Yes
1	GDP Trade, Patents	0.0197299 (0.0040629)	4.86***	86580	+	Yes
2	GDP, Trade Edu	0.0245782 (0.0031542)	7.79***	67076	+	Yes
3	GDP, Trade FDI In	0.0226147 (0.002975)	7.60***	84812	+	Yes
4	GDP, Trade FDI Out	0.0228739 (0.0029867)	7.66***	85192	+	Yes
5	GDP Patents, Edu	0.0300292 (0.0041359)	7.26***	67076	+	Yes
6	GDP, Patents FDI In	0.0254621 (0.0035345)	7.20***	84812	+	Yes
7	GDP, Patents FDI Out	0.0264495 (0.0035549)	7.44***	85192	+	Yes
8	GDP, Edu FDI In	0.0259236 (0.0030576)	8.48***	65366	+	Yes
9	GDP, Edu FDI Out	0.026174 (0.0030686)	8.53 ***	65689	+	Yes
10	GDP, FDI In FDI Out	0.0249787 (0.0028989)	8.62***	83441	+	Yes
11	Trade Patents, Edu	0.0279408 (0.0050501)	5.53***	67076	+	Yes
12	Trade, Patents FDI In	0.0201977 (0.0040871)	4.94***	84812	+	Yes
13	Trade, Patents FDI Out	0.0210359 (0.0040977)	5.13***	85192	+	Yes
14	Trade, Edu FDI In	0.0248491 (0.0031283)	7.94***	65366	+	Yes
15	Trade, Edu FDI Out	0.0249844 (0.0031493)	7.93***	65689	+	Yes
16	Trade, FDI In FDI Out	0.0234141 (0.0029872)	7.84***	83441	+	Yes
17	Patents, Edu FDI In	0.0325153 (0.0026795)	12.13***	65366	+	Yes
18	Patents, Edu FDI Out	0.032591 (0.0027058)	12.04***	65689	+	Yes
19	Patents, FDI In FDI Out	0.0327471 (0.0024828)	13.19***	83441	+	Yes
20	Edu, FDI In FDI Out	0.0328129 (0.0025038)	13.11***	64000	+	Yes

Note: Results from FE Poisson regression analysis. Standard errors are robust to heteroskedasticity
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table 7 Descriptive statistics, definitions of immigration

Immigration Variable	Obs	Mean	Std. Dev.	Min	Max
Employment-Based Immigration	237540	7.951135	10.32182	.614	48.447
ln (Employment-Based Immigration)	237540	1.467108	1.062385	-.4877603	3.88047
Total Immigration	267072	16.66218	19.43746	1.848	84.413
ln (Total Immigration)	267072	2.249879	1.043791	.614104	4.435721
Stock of Immigrants	267072	200.8276	256.6319	7.67	1204.19
ln (Stock of Immigrants)	267072	4.613896	1.162784	2.037317	7.093563
Stock of Foreign Labor	243960	134.3526	174.2607	2	633
ln (Stock of Foreign Labor)	243960	4.085989	1.345867	.6931472	6.45047

Note: For variable data sources, see Appendix

Table 8 Sensitivity analysis, definitions of immigration

Immigration Variable	Coefficient on Immigration (Standard Error)	z-statistic	Obs	Sign	Significant
Employment-Based Immigration	0.0311889 (0.00519)	6.01***	64000	+	Yes
ln (Employment-Based Immigration)	0.6160226 (0.0572536)	10.76***	64000	+	Yes
Total Immigration	0.0127475 (0.0027941)	4.56***	72900	+	Yes
ln (Total Immigration)	0.566939 (0.0622402)	9.11***	72900	+	Yes
Stock of Immigrants	0.0011552 (0.0001667)	6.93***	72900	+	Yes
ln (Stock of Immigrants)	0.5296944 (0.0469756)	11.28***	72900	+	Yes
Stock of Foreign Labor	0.0022368 (3.565e-04)	6.27***	64452	+	Yes
ln (Stock of Foreign Labor)	0.5423597 (0.0483692)	11.21***	64452	+	Yes

Note: Results from FE Poisson regression analysis with original covariates (Table 5). Standard errors are robust to heteroskedasticity
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

immigration, as opposed to employment-based immigration, also affects significantly knowledge flows. This result is somewhat surprising given that one could feasibly argue that the group of employment-based migrants would contain high-skilled migrants who would be more likely to contribute to technological advance via patent activity.

Conclusion

Technological progress has a distinct and real effect on immigration. However, this relationship is not unidirectional; it has also been shown that migration can affect technological progress, or at least the dissemination of knowledge. If knowledge flows back to source countries as a result of outward migration, they need not suffer the welfare-reducing effects associated with this migration. In fact, they could benefit from sending labor abroad if it meant expedited knowledge transfers from host countries. This process, however, remains relatively untouched in the field of economics, especially on an international scale. That is, little is known about the effects of migration on technological progress in the source country.

This paper has examined the relationship between international migration and technology flows from host to source countries in a uniquely robust way. The results have provided evidence that a positive and statistically significant relationship exists between migration flows and technology flows. This implies that migration to a host country can create knowledge or technology flows back to the source country. Furthermore, this relationship has been analyzed in an exceptionally thorough way. Sensitivity analyses were performed on the variable of immigration, showing that the relationship between immigration and knowledge flows is robust to the inclusion or omission of explanatory variables and also to different definitions of immigration. This evidence is not currently available on an international scale in any context.

The results of this paper have many implications for both source and host countries. Haque and Kim (1995) suggest that the return on human capital investment, i.e., education, can actually be negative after a certain point if it causes human capital flight. However, if some return on emigration exists in the form of increased inward technology flows, source countries may be more willing to invest in human capital. And, if immigration is creating "brain circulation" that makes both host and source countries better off, host countries, namely the US, would also benefit from relaxing stringent immigration policies.

Though the results of this paper are important, there is still much to be done in order to uncover the complete details of the correlation between technology flows and migration flows. More dependent variables can be added to the model to further isolate the effect of immigration on technology flows. In addition, it is important to determine the specific avenues through which migration enables knowledge to flow back to source countries. For example, does immigration to the US create incentives to send FDI to source countries, as Saxenian (2002) hypothesized? Is it the personal contact that immigrants maintain with residents remaining in source countries that helps facilitate knowledge transfer, as Agrawal et al. (2003) speculated? Or, is it via return migration that knowledge flows to source countries, as concluded by Mayr and Peri (2008)? Further investigation is warranted in order to answer these questions.

Finally, this paper has shown that migration creates knowledge flows from source to host countries. Though this result is significant, it was found using migration patterns between developed countries. Because the out-migration of highly educated people may be most detrimental to developing countries, this result could have an even greater impact for lesser developed countries suffering from brain drain migration. However, the data now available for patent citations does not permit us to apply the methodology used in this paper for developing economies. Clearly, this shortcoming must be addressed given that the biggest concerns about the brain drain relate to the still developing economies. Though there is still much work to be done on this topic, this paper has shed light on the possibility that immigration flows facilitate knowledge flows. It thus provides further insight into how technological advance can be achieved, technology gaps can be bridged, and sustainable long-run economic growth can be realized.

Endnotes

[1]Figure obtained from the National Science Board (2006).

[2]Information obtained from the Patent Cooperation Treaty (1970).

[3]There is a large existing literature regarding the relationship between migration and trade. There are many factors that affect both trade and migration—including but not limited to geographical distance, cultural distance and country size (Ortega and Peri 2014). These factors, however, are time invariant and cannot be included in the fixed effects model.

[4]The sample of US patents consists of only those US patents granted to inventors residing in the US. A large portion of US patents, 40 percent according to Jaffe and his colleagues, are granted to inventors residing in foreign countries.

[5]This information was gathered via the Patent Full-Text and Image Database, a database operated by the US Patent and Trademark Office (USPTO).

[6]One may note that the time period starts *before* that of the sample of US patents. This is because different dates were used to define the US patents and the citing patents; it does not alter the direction of the knowledge flow.

[7]To test whether the regression results are sensitive to model specification, the regression was run using two additional data models: FE OLS and conditional FE negative binomial. Using both regression models, immigration was found to have a positive and statistically significant effect on forward citations. Thus, the coefficient on the variable of interest is robust to regression model specification.

[8]Note that the first regression consists of immigration as the only independent variable.

Appendix

Table 9 Definitions of variables and sources of data

Variable	Source	Definition/Description
Employment-Based Immigration	US Department of Homeland Security – Annual Statistical Yearbook	Immigrants (in thousands) allowed into the US for specific, industry-based purposes, including spouses and children in given year.
Total Immigration	US Department of Homeland Security – Annual Statistical Yearbook	Immigrants (in thousands) allowed into the US in given year.
Stock of Immigrants	OECD Statistics http://stats.oecd.org/	Stock of foreign-born population (in thousands) by country of birth
Stock of Foreign Labor	OECD Statistics http://stats.oecd.org/	Stock of foreign-born labour force (in thousdands) by country of birth
GDP	World Bank http://data.worldbank.org/indicator/NY.GDP.MKTP.CD	"GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars. Dollar figures for GDP are converted from domestic currencies using single year official exchange rates. For a few countries where the official exchange rate does not reflect the rate effectively applied to actual foreign exchange transactions, an alternative conversion factor is used." Note: GDP for Australia in 2010 was missing and, but found at the IMF website: http://elibrary-data.imf.org/DataReport.aspx?c=1449311&d=33060&e=161838
Trade	U.S. Department of Commerce, Bureau of the Census, Foreign Trade http://www.census.gov/foreign-trade/balance/	Imports of US goods from sample countries plus exports to US in millions of dollars. Values are not seasonally adjusted.
Patent Stock Agriculture	Espacenet http://worldwide.espacenet.com/?locale=en_EP	Sum of total patents and patent applications from class A01 of the international patent classification (IPC) scheme. This class includes agriculture; forestry; animal husbandry; hunting; trapping; fishing.
Education	World Bank http://data.worldbank.org	Percentage of the student aged population enrolled in tertiary education in country j at time t
Inward FDI	Bureau of Economic Analysis http://www.bea.gov/iTable/index_MNC.cfm	Foreign Direct Investment in the US from abroad in millions of US Dollars, by country
Outward FDI	Bureau of Economic Analysis http://www.bea.gov/iTable/index_MNC.cfm	US Direct Investment abroad in Millions of US Dollars, by country

Competing interests

The IZA Journal of Migration is committed to the IZA Guiding Principles of Research Integrity. The author declares that she has observed these principles.

Acknowledgements

I would like to thank the editor of this journal and two additional anonymous referees for their helpful comments. I would also like to thank Hendrik van den Berg and Mary McGarvey for invaluable research assistance.

Responsible editor: Denis Fougere

Received: 22 April 2015 Accepted: 26 May 2015
Published online: 10 July 2015

References

Agrawal A, Cockburn I, McHale J (2003) Gone but Not forgotten: labor flows, knowledge spillovers and enduring social capital. NBER Working Pap No. 9950, Cambridge, MA, pp 1–35

Coe D, Helpman H (1995) International R & D spillovers. Europ Econ Rev 39:859–887

de Haas H (2006) Emerging diasporas: How governments and development agencies can support Diaspora involvement in the development of origin countries. International Migration Institute, University of Oxford

Florida R (2005) The world is spiky. Atlantic Monthly 296(3):48

Grossman M (2010) Diaspora knowledge flows in the global economy. Management Faculty Publications Paper 18, Bridgewater, MA, USA

Haque N, Kim S (1995) Human capital flight: impact of migration on income and growth. IMF Staff Papers 42(3):577–607

Hausman J, Hall B, Griliches Z (1984) Econometric models for count data with an application to the patents-R & D relationship. Econometrica 53(4):909–938

Hu A, Jaffe A (2003) Patent citations and international knowledge flow: the cases of Korea and Taiwan. Int J Ind Organ 21:849–880

Jaffe A, Trajtenberg M, Hendersen R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. Q J Econ 108(3):577–598

Kerr W (2013) U.S. High-skilled immigration, innovation, and entrepreneurship: empirical approaches and evidence. NBER Working Pap No. 19377, Cambridge, MA

Leamer A, Leonard H (1983) Reporting the fragility of regression estimates. Rev Econ Stat 65:306–317

Levine R, Renelt D (1992) A sensitivity analysis of cross-country growth regression. Am Econ Assoc 82(4):942–963

MacGarvie M (2005) The determinants of international knowledge diffusion as measured by patent citations. Econ Letters 87:121–126

Mayr K, Peri G (2008) Return migration as a channel of "Brain Gain". NBER Working Pap No. 14039, Cambridge, MA, pp 1–52

Murakami Y (2010) Japan's brain drain: an analysis of Japanese researchers living in the United States. Jpn Econ 37(2):23–57

Ortega F, Peri G (2014) The aggregate effects of trade and migration: evidence from OECD countries, The socio-economic impact of migration flows. Pop Econ. Springer International Publishing, Switzerland

Regets M (2001) Research and policy issues in high-skilled international migration: a perspective with data from the United States. IZA Discuss Pap No. 366, Bonn, Germany

Saggi K (2002) Trade, foreign direct investment, and international technology transfer: a survey. World Bank Res Obser 17(2):191–235

Saint-Paul G (2004) The brain drain: some evidence from european expatriates in the United States. IZA Discuss Pap No. 1310, Bonn, Germany

Saxenian A (2002) Brain circulation: How high-skill immigration makes everyone better Off. Brook Rev 20(1):28–31

Sjöholm F (1996) International transfer of knowledge: the role of international trade and geographic proximity. Weltwirtsch 132:97–115

Washington Diplomatic Conference on the Patent Cooperation Treaty. Patent Cooperation Treaty. June 19, 1970. Treaties and Other International Agreements