For submission to:

«The Role of Immigrants and Foreign Students in Science, Innovation, and Entrepreneurship»

Cambridge MA, April 27, 2018

Returnee migrants' self-selection: Evidence for Indian inventors in the US

Stefano Breschi[a], Francesco Lissoni[a,b], Ernest Miguelez[b,c]

[a] *ICRIOS, Università L. Bocconi, Milano*

[b] *GREThA UMR CNRS 5113, Université de Bordeaux*

[c] *AQR-IREA, Universitat de Barcelona*

EXTENDED ABSTRACT

This version: October 30, 2017- Not to be quoted nor circulated

1 Motivation and research question

Return migration represents an important share of present-day total migration flows[1]. Despite the absence of official statistics, cohort analysis of national census data suggests that, for destination countries such as Germany, the UK, and the US, the estimated rates of return after 20 years of permanence abroad vary between 20% and 80%, with much variation due to heterogeneity of sending-receiving country pairs, years of entry, and the skill level of migrants. Return rate after 5 years, which may be affected by students' migration, is significant too (Dustmann and Weiss 2007, Gundel and Peters 2008, Mayr and Peri 2008).

[1] Dustmann and Weiss (2007) define return migrants as those who settle back in their home country by their own choice, after having spent several years abroad. As far as innovation diffusion is concerned, contract migrants are important too, as they include individuals whose migration spell is exogenously determined by a working contract (as for Germany's *Guestarbeiters* in the 1960s) or a visa (as for most foreign students).

The same references we just cited suggest that the highly educated migrants' rate of return is generally higher than average, although the interaction with the conditions of the home economy is once again crucial. To put it differently, return migrants appear in general to be positively selected according to education and their home country's conditions. This basic evidence questions prevailing economic theories, as it seems at odds with classic migration models, which mostly posit skill-based negative selection, as they associate return migration to lack of economic success or integration in the host country (Borjas and Bratsberg, 1996). Evidence and models can be reconciled by observing that while positive self-selection cuts across general levels of education, negative self-selection can still occur within each category. Some evidence in this sense is provided by Gaulé (2014). We contribute to this line of research by proposing to examine a specific professional category of migrants, namely inventors, whose level of education is generally high and at the same time exhibit non-negligible rates of return.

The issue of positive vs negative self-selection, and its connection to education levels, is not only theoretically interesting, but also politically relevant, as returnee scientists, engineers, and other professionals may have a special role in the science and technology systems of their home country. In this respect, a representative study is the one by Kahn and MacGarvie (2012), who find that returnee scientists from the US have more US-based co-authors than non-migrants, thus contributing to keep their home science system in touch with the US one. Similar findings are obtained by Jonkers and Tijssen (2008) and Jonkers and Cruz-Castro (2013), respectively for samples of Chinese and Argentinean scholars, and Baruffaldi and Landoni (2012) for a survey of 497 foreign researchers in Italy and Portugal. As for Scellato et al. (2015), based on a large survey of scientific authors, they find that foreign-born in selected OECD countries have larger collaboration networks than natives; and that large networks always include contacts in both the home countries and with fellow migrants to other destination countries.

Notice however that this type of studies tend not to say much on self-selection, which they tend to consider more as a source of endogeneity problems, to be weeded out by means of quasi-experimental research designs, than a topic of interest[2].

We propose instead to focus precisely on the determinants of highly skilled migrants' return decision, based on an extensive dataset that combines inventors' data, derived from patent documentation, with curricular information extracted from LinkedIn profiles. In particular, we propose to test whether returnee inventors are positively self-selected based on two measures of skill:

1. education (in particular, post-graduate vs graduate education)
2. experience (as measured by the stock of patents filed, after controlling for age).

[2]For example, Kahn and MacGarvie (2012) concentrate on return scientists who took their PhD or postdoc in the US thanks to a Fulbright scholarship, which requires its beneficiaries to return home upon completion of their studies.

In what follows, we describe the data (section 2) and some preliminary results (section 3).

2 Data

Our database is the result of data-linkage between USPTO patent and inventor data gathered from Patentsview (`http://www.patentsview.org/web/`) and biographical information extracted from a large number of LinkedIn profiles. Patentsview is a data repository and data visualization tool recently made available by the USPTO, which provides disambiguated data on inventors of USPTO granted patents from 1975 onward. LinkedIn, a well-known social networking system, reports a very large number of users' public profiles that include information on the users' educational curricula and careers (name and possibly locations of education institutions and employer), thereby allowing to trace (return) migration with a scale and degree of precision unmatched by other sources of data. While subject to a number of limitations[3], LinkedIn represents an unparalleled source of information on the international mobility of individuals, both as students and as workers. For the purposes of the proposed paper, we extracted all the patents granted to the 178 largest US public firms in the ICT industry[4], from 1975 to 2016. For each company we identified the inventors of its patents using Patentsview, for a total of 262,849 distinct individuals.

We then proceeded to the ethnic analysis of such inventors' names and surnames, based on Global Name Recognition, a name search technology produced by IBM (from now on, IBM-GNR) and adapted to our purposes by Breschi et al. (2017). This allowed us to identify inventors of presumed Indian origin (from now on, Indian inventors), for a total of 24,017 individuals, representing 9.1% of all inventors employed by companies in our sample. Each Indian inventor was then matched to one LinkedIn profile, based on name- and company-matching, with extensive follow-up manual checking. This exercise yielded 10,839 inventors with a valid LinkedIn account (around 45% of the original Indian inventor sample). Some of the matched profiles, however, were incomplete, to the extent they did not report either information on employment or education. After dropping these cases, we were left with 7,938 inventors with a complete biography (around 33% of the original Indian inventor sample).

For each inventor, we coded two major sets of variable. One set captures the migrant and the returnee status of each inventor. To this purpose, we defined as migrant any inventor who in her career made at least one USPTO patent

[3]First, resume information is self-reported by individuals and therefore subject to misreporting or even cheating. Second, the choice of having an account in a professional network might be correlated with factors affecting the propensity to move (and migrate), thus leading to biased results.

[4]The definition of ICT industry follows the one provided by the OECD (`https://www.oecd.org/sti/ieconomy/1835738.pdf`). For each SIC code contained in the OECD definition, we extracted from Compustat the list of public US firms active in that SIC and we matched them to the USPTO patent data. Finally, we retained the firms with more than 250 patents granted.

Table 1: Migrants and non-migrants as defined on the basis of location of patenting activity

Variables	All Indian inventors	Indian inventors with a LinkedIn profile	Indian inventors with a complete LinkedIn profile
Number of early migrants	19201	8207	5879
% of all inventors	80.0	75.7	74.1
Number of late migrants	492	252	201
% of all inventors	2.0	2.3	2.5
Number of non-migrants	4324	2380	1857
% of all inventors:	18.0	22.0	23.4
Total number of inventors	24017	10839	7938

Notes: Inventors are defined as *migrants* if they made patents outside India; *early migrants* if they did not patent in India before patenting abroad; *late migrants* if they did patent in India before patenting abroad; *non-migrants* if they never made patents abroad. Inventors are considered as having a complete LinkedIn profile if the resume reports information on the employment and education career.

reporting an address outside India (not necessarily in the US). Accordingly, we assumed that the starting date of migration corresponded to the year of the first patent made abroad. These migrants can be further distinguished between early and late migrants, depending on whether they patent abroad before or after having first filed a patent to the USPTO from India. Table 1 reports some descriptive statistics.

As far as the returnee status is concerned, we provided several definitions, according to the application of more or less stringent criteria. First, we defined as patent-based returnee any inventor who made at least one patent in India after filing patents abroad. Descriptive statistics are reported in table 2.

This patent-based definition of returnee has been commonly used in the literature, but it suffers from a major problem. It does not allow to spot inventors who stop patenting at the USPTO after returning to India (either because they move out of R&D or patent only at the Indian patent office, for which we have no data). To remedy this problem, we used locational data on the inventors' employers as reported by LinkedIn profiles. More precisely, we check the LinkedIn user's workplace at the time when she declared having a job with the employer of interest.

We define as employment-based returnee any inventor who patented in India after patenting abroad and/or got a job located in India after patenting abroad. Notice however the employer's location information is often missing. In these cases, we estimated the likelihood of the location to be in India as follows. For each employer, we computed the fraction of all its employees with a LinkedIn profile (not just the inventors in our sample, but any LinkedIn profile hold-

Table 2: Migrants and non-migrants as defined on the basis of location of patenting activity

Variables	All Indian inventors	Indian inventors with a LinkedIn profile	Indian inventors with a complete LinkedIn profile
Patent based returnees	1169	573	438
% of all migrants	5.9	6.8	7.2
of which:			
Early migrants	1020	487	367
% of all returnees	87.3	85.0	83.8
Late migrants	149	86	71
% of all returnees	12.7	15.0	16.2

Notes: Inventors are defined as returnees if they made patents in India after patenting abroad; *early migrants* if they did not patent in India before patenting abroad; *late migrants* if they did patent in India before patenting abroad; *non-migrants* if they never made patents abroad. Inventors are considered as having a complete LinkedIn profile if the resume reports information on the employment and education career.

ers) who associated such employer to an Indian address. For example, 100% of LinkedIn users reported at the same time Vigyanlabs Pvt. Ltd. as an employer and an Indian workplace, which we interpret as an indication that all Indian employees of Vigyanlabs work in India, including those with a patenting experience outside India, who we then safely consider as employment-based returnees. Table 3 reports some descriptive statistics.

Based on the same principle, we can progressively extend the status of employment-based returnee to Indian inventors with patenting experience abroad and a successive employment at a company that 90%, 70%, 50% or 30% of LinkedIn users report as located. For example, a 50%- employment-based returnee will be one whose employer is located in India, according to 50% of LinkedIn profiles mentioning the same employer. Tables 4 to 6 report some descriptive statistics. As expected, the rate of return migration of inventors based on patent data is severely under-estimated. When focussing only on patent-based returnees, the fraction of migrant inventors who return is around 6%. When moving to the employment-based returnee definition, the rate of return more than doubles to around 14%, and it increases even further to around 20%-24% when using estimated employment location.

The second set of LinkedIn items that we coded concerns the education achievements of the Indian inventors. In particular, we coded the education levels attained by any inventor according to the ISCED definition[5]: PhD, Master of Science (MSC), Master in Business Administration (MBA), Bachelor of Science (BSC) and other unclassified titles (OTHER). Moreover, for each education level, we were able to establish whether the corresponding degree title was obtained in the home country (India) or in a foreign one.

[5] http://uis.unesco.org/en/topic/international-standard-classification-education-isced

Table 3: Returnees defined on the basis of location of patenting activity and employment (whole sample)

			Estimated location				
Variables	Patent only	Self-reported	100%	90%	70%	50%	30%
Patent or employment based returnees	573	1179	1329	1410	1576	1723	2030
% of all migrants (patent abroad)	6.8	13.9	15.7	16.7	18.6	20.4	24.0
of which:							
Employment but not patent based returnees		606	756	837	1003	1150	1457
% of all migrants (patent abroad)		7.2	8.9	9.9	11.9	13.6	17.2
Patent but not employment based returnees	573	348	315	304	291	280	262
% of all migrants (patent abroad)	6.8	4.1	3.7	3.6	3.4	3.3	3.1
Patent and employment based returnees		225	258	269	282	293	311
% of all migrants (patent abroad)		2.7	3.1	3.2	3.3	3.5	3.7

Notes. Inventors are defined as migrants if they made patents outside India. The table is based on all Indian inventors with a LinkedIn profile.

Table 4: Returnees defined on estimated employment location (whole sample)

		Estimated location				
Variables	Self-reported	100%	90%	70%	50%	30%
Employment based returnees	831	1014	1106	1285	1443	1768
of which:						
estimated	-	183	275	454	612	937
self-reported	831	831	831	831	831	831
% of all migrants	9.9	12.0	13.1	15.2	17.1	20.9

Notes. Inventors are defined as *migrants* if they made patents outside India; as returnees if they reported an employer located in India after patenting abroad. The table is based on all Indian inventors with a LinkedIn profile.

Table 5: Returnees defined on the basis of location of patenting activity and employment (reduced sample)

			Estimated location				
Variables	Patent only	Self-reported	100%	90%	70%	50%	30%
Patent or employment based returnees	438	887	996	1056	1172	1274	1499
% of all migrants (patent abroad)	7.2	14.6	16.4	17.4	19.3	21	24.7
of which:							
Employment *but not* patent based returnees		449	558	618	734	836	1061
% of all migrants (patent abroad)		7.4	9.2	10.2	12.1	13.7	17.4
Patent *but not* employment based returnees	438	257	226	217	209	200	189
% of all migrants (patent abroad)	7.2	4.2	3.7	3.6	3.4	3.3	3.1
Patent *and* employment based returnees		181	212	221	229	238	249
% of all migrants (patent abroad)		3.0	3.5	3.6	3.8	3.9	4.1

Notes. Inventors are defined as migrants if they made patents outside India. The table is based on the reduced sample of Indian inventors with a complete LinkedIn profile. Inventors are considered as having a complete LinkedIn profile if the resume reports information on the employment and education career.

Table 6: Returnees defined on estimated employment location (reduced sample)

		Estimated location				
Variables	Self-reported	100%	90%	70%	50%	30%
Employment based returnees	630	770	839	963	1074	1310
of which:						
estimated	-	140	209	333	444	680
self-reported	630	630	630	630	630	630
% of all migrants	10.4	12.7	13.8	15.8	17.7	21.5

Notes. Inventors are defined as *migrants* if they made patents outside India; as returnees if they reported an employer located in India after patenting abroad. The table is based on the reduced sample of Indian inventors with a complete LinkedIn profile. Inventors are considered as having a complete LinkedIn profile if the resume reports information on the employment and education career.

3 Preliminary results

Our preliminary results are based on a series of z-tests of proportions (tables 7a-g) and a set of discrete-time duration analysis regressions (tables 8 and 9). They can be summarized as follows:

- Migrants are found to be positively selected (i.e. non-migrants are less likely to have a higher degree, such as PhD or MSC than migrants), while returnees are found to be negatively selected (i.e. they are less likely to have higher degree, such as PhD or MSC than permanent migrants, though more likely to have a MBA). These results are robust to different definitions of returnees and are consistent with the conjecture of Borjas and Bratsberg (1996).

- A survival (duration) regression analysis shows that the likelihood of returning depends on the channel through which inventors migrated (i.e. patented) in the United States. Inventors who were either employed in India or patented in India before patenting abroad are much more likely to return. Moreover, the place where higher education has been achieved also matters. Inventors who got their PhD or MSC in India are much more likely to return compared to inventors who went abroad to get their higher education.

Table 7a: Migrant selection based on highest education achievement, Patent based returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25.0	50.4	4.6	15.4	4.4
Permanent migrants	1409	2887	256	837	248
% of permanent migrants	25.0	51.2	4.5	14.8	4.4
Returnees	112	180	25	99	22
% of returnees	25.6	41.1	5.7	22.6	5.0
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.201 (0.000)	-6.243 (0.000)	-0.218 (0.828)	22.018 (0.000)	3.842 (0.000)
z-test: non-migrants vs. returnees	-10.961 (0.000)	0.653 (0.514)	-1.154 (0.249)	6.335 (0.000)	1.239 (0.215)
z-test: permanent migrants vs. returnees	-0.280 (0.779)	4.058 (0.000)	-1.125 (0.261)	-4.341 (0.000)	-0.615 (0.539)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 7b: Migrant selection based on highest education achievement, Patent based plus self-reported returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25	50.4	4.6	15.4	4.4
Permanent migrants	1347	2662	224	731	224
% of permanent migrants	25.9	51.3	4.3	14.1	4.3
Returnees	174	405	57	205	46
% of returnees	19.6	45.7	6.4	23.1	5.2
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.783 (0.000)	-6.245 (0.000)	0.187 (0.852)	22.552 (0.000)	3.951 (0.000)
z-test: non-migrants vs. returnees	-9.405 (0.000)	-1.407 (0.159)	-2.246 (0.025)	8.091 (0.000)	1.465 (0.143)
z-test: permanent migrants vs. returnees	4.015 (0.000)	3.078 (0.002)	-2.771 (0.006)	-6.893 (0.000)	-1.167 (0.243)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad and/or reported an employment in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India *and* did not report an employment in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 7c: Migrant selection based on highest education achievement, Patent based plus self-reported and estimated (100%) returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25	50.4	4.6	15.4	4.4
Permanent migrants	1326	2605	217	711	220
% of permanent migrants	26.1	51.2	4.3	14	4.3
Returnees	195	462	64	225	50
% of returnees	19.6	46.4	6.4	22.6	5
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.857 (0.000)	-6.211 (0.000)	0.269 (0.788)	22.558 (0.000)	3.909 (0.000)
z-test: non-migrants vs. returnees	-9.632 (0.000)	-1.833 (0.067)	-2.323 (0.020)	8.727 (0.000)	1.711 (0.087)
z-test: permanent migrants vs. returnees	4.330 (0.000)	2.796 (0.005)	-2.967 (0.003)	-6.884 (0.000)	-0.972 (0.331)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad and/or reported an employment (either self-reported or estimated) in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India *and* did not report an employment (either self-reported or estimated) in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 7d: Migrant selection based on highest education achievement, Patent based plus self-reported and estimated (90%) returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25	50.4	4.6	15.4	4.4
Permanent migrants	1313	2577	211	698	220
% of permanent migrants	26.1	51.3	4.2	13.9	4.4
Returnees	208	490	70	238	50
% of returnees	19.7	46.4	6.6	22.5	4.7
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.879 (0.000)	-6.241 (0.000)	0.395 (0.693)	22.609 (0.000)	3.800 (0.000)
z-test: non-migrants vs. returnees	-9.837 (0.000)	-1.876 (0.061)	-2.582 (0.010)	8.938 (0.000)	2.073 (0.038)
z-test: permanent migrants vs. returnees	4.387 (0.000)	2.884 (0.004)	-3.419 (0.001)	-7.078 (0.000)	-0.512 (0.609)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad and/or reported an employment (either self-reported or estimated) in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India *and* did not report an employment (either self-reported or estimated) in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 7e: Migrant selection based on highest education achievement, Patent based plus self-reported and estimated (70%) returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25	50.4	4.6	15.4	4.4
Permanent migrants	1288	2525	202	673	215
% of permanent migrants	26.2	51.4	4.1	13.7	4.4
Returnees	233	542	79	263	55
% of returnees	19.9	46.2	6.7	22.4	4.7
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.925 (0.000)	-6.333 (0.000)	0.551 (0.582)	22.709 (0.000)	3.781 (0.000)
z-test: non-migrants vs. returnees	-10.179 (0.000)	-1.854 (0.064)	-2.778 (0.005)	9.322 (0.000)	2.201 (0.028)
z-test: permanent migrants vs. returnees	4.515 (0.000)	3.193 (0.001)	-3.847 (0.000)	-7.442 (0.000)	-0.468 (0.640)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad and/or reported an employment (either self-reported or estimated) in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India *and* did not report an employment (either self-reported or estimated) in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 7f: Migrant selection based on highest education achievement, Patent based plus self-reported and estimated (50%) returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25	50.4	4.6	15.4	4.4
Permanent migrants	1261	2465	197	664	214
% of permanent migrants	26.2	51.3	4.1	13.8	4.5
Returnees	260	602	84	272	56
% of returnees	20.4	47.3	6.6	21.4	4.4
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.896 (0.000)	-6.200 (0.000)	0.580 (0.562)	22.477 (0.000)	3.627 (0.000)
z-test: non-migrants vs. returnees	-10.709 (0.000)	-2.456 (0.014)	-2.672 (0.008)	10.265 (0.000)	2.638 (0.008)
z-test: permanent migrants vs. returnees	4.268 (0.000)	2.556 (0.011)	-3.772 (0.000)	-6.628 (0.000)	0.087 (0.931)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad and/or reported an employment (either self-reported or estimated) in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India *and* did not report an employment (either self-reported or estimated) in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 7g: Migrant selection based on highest education achievement, Patent based plus self-reported and estimated (30%) returnees

Variables	PHD	MSC	MBA	BSC	OTHER
Non migrants	138	795	82	719	123
% of non-migrants	7.4	42.8	4.4	38.7	6.6
Migrants	1521	3067	281	936	270
% of migrants	25	50.4	4.6	15.4	4.4
Permanent migrants	1210	2349	180	629	208
% of permanent migrants	26.4	51.3	3.9	13.7	4.5
Returnees	311	718	101	307	62
% of returnees	20.7	47.9	6.7	20.5	4.1
z-test: non-migrants vs. migrant	-16.309 (0.000)	-5.754 (0.000)	-0.371 (0.711)	21.658 (0.000)	3.796 (0.000)
z-test: non-migrants vs. permanent migrants	-16.955 (0.000)	-6.149 (0.000)	0.897 (0.370)	22.329 (0.000)	3.431 (0.001)
z-test: non-migrants vs. returnees	-11.265 (0.000)	-2.945 (0.003)	-2.945 (0.003)	11.401 (0.000)	3.139 (0.002)
z-test: permanent migrants vs. returnees	4.393 (0.000)	2.263 (0.024)	-4.497 (0.000)	-6.289 (0.000)	0.658 (0.510)

Notes. Inventors are defined as *Returnees* if they made patents in India after patenting abroad and/or reported an employment (either self-reported or estimated) in India after patenting abroad. *Non-migrants* represent inventors that never made patents outside India. *Permanent migrants* are defined as those inventors that did not patent in India *and* did not report an employment (either self-reported or estimated) in India after patenting abroad. PHD indicates doctoral degree, MSC master of science, MBA master in business administration, BSC bachelor of science, OTHER not classified titles. For each education level and each pair of inventor types, we test the difference in the proportion of individuals whose highest education achievement corresponds to that level by means of a z-test; p-values are reported in parenthesis.

Table 8: Duration analysis: time to return for various definitions of returnees (clog-log estimates)

			Estimated location				
Variables	Patent-based	Self-reported	100%	90%	70%	50%	30%
Log time clock	-0.145	-0.072	-0.068	-0.067	-0.092	-0.108*	-0.219***
	(0.087)	(0.058)	(0.054)	(0.053)	(0.050)	(0.047)	(0.042)
Profile reports knowledge of english	0.015	0.241	0.183	0.198	0.134	0.120	0.151
	(0.168)	(0.124)	(0.113)	(0.110)	(0.101)	(0.096)	(0.089)
Estimated age in year	-0.043	-0.141***	-0.128***	-0.126***	-0.113***	-0.103***	-0.071***
	(0.052)	(0.021)	(0.018)	(0.016)	(0.015)	(0.015)	(0.016)
Estimated age squared	-0.000	0.001***	0.001***	0.001***	0.001***	0.001***	0.001**
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log of cumulative patents	0.700***	0.313***	0.251***	0.210***	0.193***	0.164***	0.143***
	(0.048)	(0.036)	(0.034)	(0.033)	(0.032)	(0.031)	(0.029)
Inventor patent in India before patenting abroad	1.774***	1.585***	1.568***	1.539***	1.459***	1.397***	1.273***
	(0.142)	(0.112)	(0.110)	(0.109)	(0.107)	(0.107)	(0.103)
Inventor employed in India before patenting abroad	0.689***	1.375***	1.262***	1.218***	1.104***	1.044***	0.930***
	(0.130)	(0.080)	(0.077)	(0.076)	(0.074)	(0.072)	(0.068)
0/1 if PHD highest degree in year	-0.540***	-0.654***	-0.645***	-0.627***	-0.621***	-0.519***	-0.488***
	(0.140)	(0.105)	(0.099)	(0.096)	(0.091)	(0.087)	(0.080)
0/1 if MBA highest degree in year	-0.382	-0.034	-0.015	0.050	0.058	0.166	0.177
	(0.211)	(0.129)	(0.120)	(0.113)	(0.107)	(0.101)	(0.094)
0/1 if MSC highest degree in year	-0.373**	-0.419***	-0.412***	-0.444***	-0.441***	-0.370***	-0.349***
	(0.117)	(0.080)	(0.076)	(0.074)	(0.070)	(0.068)	(0.063)
0/1 if BSC highest degree in year		0.328	0.214	0.136	0.028	-0.003	-0.126
		(0.711)	(0.710)	(0.710)	(0.710)	(0.710)	(0.709)
Observations	76706	73756	72829	72298	71417	70565	68744
Log-likelihood	-2437.0	-4417.3	-4907.8	-5162.3	-5650.0	-6072.2	-6915.6
Chi2	524.6	771.0	712.5	697.8	658.0	609.1	603.0

Notes: The table is based on the reduced sample of Indian inventors with a complete LinkedIn profile. Inventors are considered as having a complete LinkedIn profile if the resume reports information on the employment and education career. * significant at 10%, ** significant at 5%, *** significant at 1%. Standard errors in parenthesis.

Table 9: Duration analysis: time to return for various definitions of returnees (clog-log estimates)

			Estimated location				
Variables	Patent-based	Self-reported	100%	90%	70%	50%	30%
Log time clock	-0.074	0.002	0.007	0.010	-0.011	-0.036	-0.165***
	(0.086)	(0.058)	(0.055)	(0.053)	(0.050)	(0.048)	(0.043)
Profile reports knowledge of english	-0.066	0.186	0.135	0.152	0.089	0.084	0.125
	(0.169)	(0.124)	(0.113)	(0.110)	(0.101)	(0.096)	(0.089)
Estimated age in year	-0.046	-0.152***	-0.142***	-0.141***	-0.128***	-0.118***	-0.080***
	(0.052)	(0.023)	(0.019)	(0.017)	(0.016)	(0.016)	(0.017)
Estimated age squared	-0.000	0.001***	0.001***	0.001***	0.001***	0.001***	0.001**
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log of cumulative patents	0.696***	0.300***	0.238***	0.197***	0.182***	0.154***	0.134***
	(0.048)	(0.036)	(0.034)	(0.033)	(0.032)	(0.031)	(0.028)
Inventor patent in India before patenting abroad	1.603***	1.419***	1.417***	1.378***	1.284***	1.236***	1.142***
	(0.143)	(0.114)	(0.111)	(0.110)	(0.109)	(0.108)	(0.104)
Inventor employed in India before patenting abroad	0.523***	1.272***	1.159***	1.114***	0.998***	0.948***	0.853***
	(0.132)	(0.082)	(0.079)	(0.077)	(0.075)	(0.073)	(0.069)
0/1 if PHD highest degree in year	-0.611***	-0.712***	-0.689***	-0.684***	-0.675***	-0.560***	-0.521***
	(0.144)	(0.108)	(0.102)	(0.099)	(0.093)	(0.089)	(0.081)
PHD × achieved in India	0.935**	0.974***	0.831**	0.969***	0.937***	0.789**	0.664**
	(0.360)	(0.277)	(0.274)	(0.252)	(0.245)	(0.243)	(0.230)
0/1 if MBA highest degree in year	-0.921**	-0.460**	-0.396**	-0.297*	-0.287*	-0.108	-0.036
	(0.284)	(0.160)	(0.145)	(0.135)	(0.126)	(0.116)	(0.106)
MBA *× achieved in India	1.585***	1.675***	1.599***	1.571***	1.652***	1.446***	1.214***
	(0.381)	(0.230)	(0.215)	(0.205)	(0.194)	(0.188)	(0.181)
0/1 if MSC highest degree in year	-0.844***	-0.721***	-0.710***	-0.737***	-0.733***	-0.627***	-0.535***
	(0.140)	(0.092)	(0.087)	(0.084)	(0.080)	(0.077)	(0.070)
MSC× achieved in India	1.354***	0.994***	1.002***	0.997***	1.007***	0.920***	0.705***
	(0.160)	(0.113)	(0.107)	(0.105)	(0.100)	(0.096)	(0.090)
0/1 if BSC highest degree in year		-11.701	-13.243	-12.293	-12.041	-12.612	-12.980
		(468.121)	(937.315)	(550.848)	(459.973)	(603.964)	(706.852)
BSC× achieved in India		12.351	13.802	12.789	12.426	12.984	13.230
		(468.122)	(937.315)	(550.849)	(459.973)	(603.964)	(706.853)
Observations	76706	73756	72829	72298	71417	70565	68744
Log-likelihood	-2394.1	-4355.5	-4841.6	-5092.4	-5570.3	-6003.3	-6866.0
Chi2	610.3	894.5	845.0	837.6	817.5	747.0	702.1

Notes: The table is based on the reduced sample of Indian inventors with a complete LinkedIn profile. Inventors are considered as having a complete LinkedIn profile if the resume reports information on the employment and education career. * significant at 10%, ** significant at 5%, *** significant at 1%. Standard errors in parenthesis.

References

Baruffaldi, S. H. and Landoni, P. (2012). Return mobility and scientific productivity of researchers working abroad: The role of home country linkages. *Research Policy*, 41(9):1655–1665.

Borjas, G. J. and Bratsberg, B. (1996). Who Leaves? The Outmigration of the Foreign-Born. *The Review of Economics and Statistics*, 78(1):165.

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

Dustmann, C. and Weiss, Y. (2007). Return Migration: Theory and Empirical Evidence from the UK. *British Journal of Industrial Relations*, 45(2):236–256.

Gaulé, P. (2014). Who comes back and when? Return migration decisions of academic scientists. *Economics Letters*, 124(3):461–464.

Gundel, S. and Peters, H. (2008). What determines the duration of stay of immigrants in Germany? *International Journal of Social Economics*, 35(11):769–782.

Jonkers, K. and Cruz-Castro, L. (2013). Research upon return: The effect of international mobility on scientific ties, production and impact. *Research Policy*, 42(8):1366–1377.

Jonkers, K. and Tijssen, R. (2008). Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity. *Scientometrics*, 77(2):309–333.

Kahn, S. and MacGarvie, M. (2012). The Effects of the Foreign Fulbright Program on Knowledge Creation in Science and Engineering. In Lerner, J. and Stern, S., editors, *The Rate and Direction of Inventive Activity Revisited*, pages 161–197. University of Chicago Press.

Mayr, K. and Peri, G. (2008). Return Migration as a Channel of Brain Gain. Technical report, National Bureau of Economic Research, Cambridge, MA.

Scellato, G., Franzoni, C., and Stephan, P. (2015). Migrant scientists and international networks. *Research Policy*, 44(1):108–120.