WHERE DO IMMIGRANT SCIENTISTS WORK?

EVIDENCE FROM ARTIFICIAL INTELLIGENCE PHDS

Caroline Fry[a], Britta Glennoni[b]

[a]Shidler College of Business, University of Hawaii at Manoa

[b]The Wharton School of the University of Pennsylvania and NBER

December 28, 2022

Abstract

Graduates of STEM PhD programs frequently make the choice of working either in industry or in academia, a choice which shapes the broader direction of science and innovation. Recent research has considered a range of factors influencing this choice, including preferences and ability. In this paper we explore the role of immigrant status. Using data on the career history of more than 1,750 graduates of U.S. PhD programs whose dissertations pertained to Artificial Intelligence (AI) we assess the extent to which immigrants and natives differ in their career choices following graduation. We document that immigrant PhDs are significantly more likely to work in industry as compared to academia, relative to non-immigrant PhDs, and we explore potential mechanisms. Our findings have implications for research on immigrant labor markets, studies of occupational choice, and diaspora knowledge flows.

Keywords: Immigration; Science; Knowledge Diffusion; Artificial Intelligence; Early Career

1 Introduction

STEM PhD graduates are especially important to the US economy; they are disproportionate generators of basic research as represented by publications and traditionally produced in academic institutions, and, increasingly also of applied research, as represented by patents and traditionally produced in the private sector. These two paths also have broad implications for knowledge flows; the academic tradition is to disseminate research results to the broader academic community around the globe, while the private sector operates under a closed commercial logic. PhD graduates' choice of industry or academia thus shapes the broader direction of science and innovation.

Earlier work has considered a range of supply-side factors that might influence the industry-academia choice, such as ability a taste for science, and preferences for money and resources (e.g. Stern 2004; Gans et al 2010; Roach and Sauermann 2010; Sauermann and Roach 2014; Agarwal and Ohyama 2013), as well as variation on the demand side for scientists across institution types (Agarwal and Ohyama 2013). Yet much of this work views scientists as a homogenous group, or, if considering heterogeneous features, examine variation only along ability and preferences (Roach and Sauermann 2010).

But increasingly, PhD programs–and particularly science and engineering PhD programs–include large numbers of immigrant doctorates; indeed, in 2020, temporary visa holders accounted for 34 percent of all doctorate degrees and in fact exceeded the number of US citizen doctorate holders in STEM fields, with these proportions rising over time (National Science Foundation). Immigrant doctorates might look quite different from native doctorates for three main reasons. First, they might bring a different skill or ability set than natives. Natives may have a comparative advantage in language-intensive type positions (such as teaching- and presentation-intensive faculty positions) while immigrants might have a comparative advantage in technical, research-intensive tasks (e.g. Peri and Sparber 2009). Immigrants might be unusually talented; numerous frictions including search costs, discrimination, and more might prevent immigrants from entry into a PhD program, so any immigrants able to overcome these frictions might be especially able or motivated (Roy 1951; Borjas 1986; Hernandez and Kulchina 2020; Wang 2015; Glennon et al 2021). Due to the tacitness and stickiness of knowledge, immigrant PhDs might also bring novel, different ideas and routines to their organizations, which can be newly utilized and/or recombined (e.g. Choudhury and Kim 2019, Moser et al 2014; Moser and San 2020; Bernstein et al 2019). Second, immigrants might have different preferences than natives. Perceptions of prestige, financial needs (e.g. Clemens 2013), preferences for stability, and preferences for proximity to diasporas are likely to vary across the two groups, affecting the industry-academia choice. And finally, institutional constraints–like the likelihood of receiving a visa

or permanent residency–could also impact the occupational choice. The Optional Practical Training (OPT) program allows STEM PhDs to work in the US for up to three years, but immigrant PhD s must obtain an employment-based temporary work visa–such as the H-1B visa–to remain longer in the private sector. The private sector is subject to a stringent cap on H-1B visas, while universities have no such constraint. Thus, immigrant PhDs might perceive that tenure-track faculty positions are more stable due to the increased likelihood of getting a work visa.

In this paper, we investigate the occupational choices of recent graduates from U.S. PhD programs, and more specifically, how they differ for immigrant and native PhD program graduates. Our base sample is the Center for Security and Emerging Technology (CSET) database of the career history of more than 1,750 graduates of U.S. PhD programs whose dissertations pertained to Artificial Intelligence (AI). We combine career data with publication and patent data for each scientist using OpenAlex publication database and PatentsView, respectively. These data allow us to examine the career paths of AI doctorates, both before and after they obtain their PhD, in terms of their occupational choices, productivity outcomes as measured using patents and publications, and implications for cross-border knowledge flows. The focus on AI is useful for two main reasons. First, AI has exceptional potential for altering the innovative landscape, the broader economy, and society at large. Advances in AI not only have the potential to directly introduce a wide range of new products and services, but they also could change the innovation process itself (Cockburn, Henderson, and Stern 2019). Second, because AI encompasses many domains, from healthcare to engineering, our focus on doctorates with AI-relevant dissertations allows us to tease out supply-side and demand-side factors; two graduates who study AI but who graduate in different fields will face very different demand characteristics–such as time to tenure or compensation–from universities and private sector companies.

We find that immigrant PhDs are significantly more likely to work in industry as compared to academia, relative to non-immigrant PhDs. We find that the results are particularly salient for Chinese immigrants and reflect the choice to work at multinational corporations rather than smaller companies. We consider three main perspectives in explaining this finding: (1) skill- and ability-based differences, (2) preference-based differences, and (3) visa availability-based differences. Our results have important implications for the scientific landscape given the critical role played by academia in higher education and innovation. As the composition of STEM PhD programs continues to change towards more and more foreign students, understanding how such foreign students might make different occupational choices than native students will help us to understand the resultant changing workforce composition and its implications for the balance between applied and basic research and the path of knowledge flows. Our findings also contribute to research on immigrant labor markets, studies of occupational choice, and diaspora knowledge flows.

2 Data

2.1 Sample of Artificial Intelligence PhD Graduates

The scientists in our sample are those who graduate from U.S. PhD programs in Artificial Intelligence related fields between 2014 and 2018 inclusive. Our career data comes from the Center for Security and Emerging Technology (CSET). CSET collected comprehensive education (undergraduate and any post graduate information, including information on institution, department, advisor and date of graduation), and career data on more than 1,750 PhD graduates from the 20 highest-ranking U.S. programs for AI-related fields. CSET generated the database by searching ProQuest for relevant graduates[1] and then manually tried to find each PhD recipient online. The employment record of each graduate for up to 6 years after completion of their the PhD is documented. We supplement missing data by manual internet searches. We use data on their career choices (employer, job position and dates of employment) in the first few years after graduating from the PhD program. Employers are manually categorized into three groups: academia, industry and government/non-profit.

As we do not have access to the country of birth of the sample scientists, we use a variety of proxies for immigration status. Our primary indicator of whether or not a scientist is an immigrant is whether or not they completed their undergraduate degree outside of the U.S. This measure provides us with a conservative estimate; it will not include those who immigrated to the US at a young age or for college, but only about three percent of undergraduates are immigrant students (Bound et al 2015). Alternatively, to provide an upper bound, we assign name ethnicity to each sample scientist using the NamePrism algorithm. This measure provides bias in the opposite direction in that it will include many natives who are descendants of immigrants. Between the two, however, we are confident that we can triangulate our results.

Our main sample consists of 1,769 AI doctorates. Just over half of the PhD graduates in our sample are immigrants, as defined using the undergraduate degree definition, with Chinese immigrants making up the largest portion at 20 percent of the full sample and Indian immigrants making up the second largest portion at 8 percent of the full sample. Immigrants from high-income countries (defined as OECD countries) comprise 7 percent of the full sample. In A2 we provide details on the set of institutions and formal departments that the individuals in our sample graduated from.

2 shows differences between native and immigrant scientists in our sample across several characteristics. While there is no statistical difference in graduation year, there are large statistical differences in pre-PhD

[1] A comprehensive query of full text of all PhD dissertations was used to identify AI related theses. Some example search phrases include "information retrieval", "autonomous vehicles" and "random forest". The resulting theses were then filtered for false-positive results (dissertations using keywords in a context unrelated to AI and machine learning).

productivity measures (publications and patents) for the two groups, where immigrant students appear to be much more productive prior to the PhD, potentially supporting the idea that they are, indeed more talented. We will test this more formally later.

A final important statistic to note is that more than 90 percent of PhD graduates in the sample continue to work in the U.S. following graduation, with no statistical difference between native and immigrant PhD graduates. These numbers are broadly consistent with prior work measuring stay rates of immigrant graduates. For example, Finn (2012) finds that 89% and 79% of PhD holders in STEM fields from China and India remain in the US five years after receiving their PhD.

2.2 Publication and Patent Data

Each scientist in our sample is matched to their full publication history using the OpenAlex API. One of the main challenges associated with using publication databases is the common name problem. Fortunately, OpenAlex generates scientist identifiers and clusters together publications of scientists based on topic, collaboration patterns and institution. We use this identifier to extract publications for our sample scientists by matching names and known affiliated institutions. For some scientists we have several OpenAlex identifiers returned who have the same name and are affiliated with the same institution in the same time period. In these instances we use data on the scientist's advisor and department to extract the best match (or matches). We run a very similar procedure to collect patent data for each scientist in the sample using the PatentsView API.

2.3 NSF Survey of Doctorate Recipients

In order to assess some potential mechanisms we supplement the data on the sample of AI PhD graduates with data from the NSF Survey of Doctorate Recipients. We extract from this dataset the full sample of science and engineering students who graduate between 2010 and 2015 who responded to the survey. The survey asks graduates from U.S. PhD programs about their career choices, salary, expectations and motivations. It also collects information on their visa status at the time of graduation.

3 Results

3.1 Occupational Choice

1 illustrates raw differences in occupational choice for immigrant and non-immigrant PhDs. It clearly illustrates, at least in the raw data, that immigrant PhDs are significantly more likely to work in the private

sector after graduating than non-immigrant PhDs. But there could be significant omitted variable bias here; perhaps, for example, immigrant students are more likely to attend PhD programs at institutions that funnel graduates to the private sector. In the remainder of this section, we therefore analyse this choice in a regression framework to account for alternative observable factors that could be driving this trend.

$$Y_{ij} = \beta_0 + \beta_1 Immigrant_i + \beta_2 PhDpublications + \delta_{graduationyear} + \gamma_{institution} + \sigma_{department} \quad (1)$$

Where *Y* is a the outcome measure for person i, and in the majority of specifications for job j. Immigrant denotes an indicator variable that switches to one the graduate received their undergraduate degree from a non-US institution. $\delta_{graduationyear}$ stands for a full set of graduation calendar year indicator variables to account for the fact that aggregate research activities may vary over time. $\gamma_{institution}$ correspond to institution fixed effects and $\sigma_{department}$ correspond to department fixed effects. Standard errors are clustered at the level of the institution of the advisor.

3 illustrates remarkable robustness of the core result to an array of fixed effects and controls for the department type, institution, graduation year, and number of authored publications in the 5 years prior to graduation. It shows that immigrant AI graduates were 10% more likely to select into industry than non-immigrants. The trend is more salient for the second job out of PhDs as compared to first job out of a PhD (columns 5 and 6)[2]. 4 utilizes an alternative measure for immigrant PhDs–non-European name ethnicity– and finds the same result; that immigrant AI doctorates are more likely to work in industry after graduation.

We explore the dynamics of these effects in 2, where panel a plots the relative propensity to be in an academic job after graduating while panel b plots the relative propensity to be in an industry job after graduation for immigrants and non-immigrants. The difference is again remarkably consistent over time, with perhaps a slight relative increase in the propensity to be in industry for immigrant graduates.

In 5, we take a deeper look to see whether there are particular roles or companies driving the stylized fact. The tables show that immigrant PhDs are much more likely to be working for a multinational corporation and in a research role in industry than non-immigrant PhDs. The immigrant preference for MNCs over other types of private companies could be driven by (1) the ease of obtaining a visa (Roach Skrentny 2018), (2) the preference for the ability to move to other countries within the same company, or (3) a preference for stability and/or prestige. We intend to investigate these possibilities further in the mechanisms section. We also explore which non-US undergraduates are driving the result in 6. We document that the trend is much

[2]The number of observations is smaller because not every person in our sample has a second job

more salient for Chinese immigrants, and is driven exclusively by low-income country immigrants.

3.2 Mechanisms

INCOMPLETE DRAFT SECTION

We consider a number of plausible mechanisms that could be driving the observed effects. One possibility is that immigrant PhDs have different opportunities available to them, either because of quality, type of research, language or visa availability. Alternatively, immigrant PhDs could have different preferences in terms of salary, job security, intellectual freedom and career advancement.

First, we consider whether differences in quality or applied nature of dissertation research affect occupational choice and thus could be driving our results. In addition to including fixed effects for department and institution and controlling for pre-PhD graduation volume of publications, we consider the role of imprinting from the advisor. In 7 we include the number of patents filed by their advisor in column 2. The results remain robust to the inclusion of this control variable. We also interact a number of measures of quality, such as publications and patents, and whether or not their university is a top ten research institute with the immigrant undergraduate variable to assess the role of these measures in driving the result. We fail to find any evidence that the occupational choice of immigrant PhDs is driven by these measures.

Second, we consider whether differences in preferences could explain the observed variation in occupational choice. Using the NSF survey data we compare preferences of recent PhD graduates in 4. When asked if they consider a factor very important in selecting a job, immigrant PhDs were significantly more likely to place value on career advancement possibilities, responsibility and salary. They were less likely to place value on independence and job location. We explore further the role of career advancement and salary differences between industry and academia. Given that AI is a broad term that captures several disciplines and university departments, we explore whether variation across PhD department in the rate of promotion in academia and the salary gap between industry and academia mirrors the occupational choice across departments. Insofar as immigrant PhDs value career advancement and financial reward following their PhD to a greater extent than non-immigrant PhDs, we should expect to see greater industry entry by immigrant PhDs in disciplines in which there is both more delayed tenure and a greater salary gap. As can be seen in 6, the fastest tenure rates in academia are in computer sciences, social sciences and biological sciences. And in 7. we show that the greatest salary gap is in computer and mathematical sciences and physical sciences, while the smallest gap is in biological and social sciences. In 8 we interact the immigrant indicator with each department indicator and find that indeed it is in computer science departments that we see the greatest gap

between immigrant and non-immigrant PhDs in industry employment post-PhD (column 6). This implies that the salary gap can explain at least part of the occupational choice gap.

3.3 Consequences of Occupational Choice

Lastly, we explore the implications of occupational choice on post-graduation production of knowledge. Specifically we explore whether immigrant PhDs have more or fewer publications and patents than non-immigrant PhDs following graduation, and whether any differences are accounted for by their choice of employer. 9 shows that immigrant PhDs are no more or less likely to publish or patent in a given year following PhD graduation (columns 1 and 5). Unsurprisingly, scientists who go to an industry job publish less and patent more than those who go to universities or government (columns 2 and 6). However, interestingly, immigrants in an industry job patent more than non-immigrants in an industry job (column 8) although they do not publish any more prolifically (column 4).

In future iterations, we hope to also explore the implications for knowledge flows more generally; immigrant scientists have connections at home and thus may be better at transferring knowledge back home, but the choice of industry over academia may also reduce knowledge flows generally.

4 Discussion

In this paper we document a large gap in the employment outcome of immigrant and non-immigrant AI PhDs. Immigrant PhDs are significantly more likely to work in industry in their first few years out of PhD. This has implications for our understanding of their contribution to the economy and to knowledge diffusion patterns, particularly diffusion of knowledge back home. Given the rising number of immigrants studying in the U.S., a better understanding of the frictions they face post graduation and how this affects their choice of workplace is critical.

References

Agarwal, Rajshree, and Atsushi Ohyama. ”Industry or academia, basic or applied? Career choices and earnings trajectories of scientists.” Management Science 59.4 (2013): 950-970.

Bernstein, Shai, et al. ”The contribution of high-skilled immigrants to innovation in the United States.” Standford Graduate School of Business Working Paper 3748 (2018): 202019-20.

Borjas, George J. ”The self-employment experience of immigrants.” (1986).

Bound, John, et al. ”Recruitment of foreigners in the market for computer scientists in the United States.” Journal of labor economics 33.S1 (2015): S187-S223.

Clemens, Michael A. ”What do we know about skilled migration and development.” Policy Brief 3 (2013).

Choudhury, Prithwiraj, and Do Yoon Kim. ”The ethnic migrant inventor effect: Codification and recombination of knowledge across borders.” Strategic Management Journal 40.2 (2019): 203-229.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern. ”4. The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis.” The Economics of Artificial Intelligence. University of Chicago Press, 2019. 115-148.

Gans, Joshua S., and Scott Stern. ”Is there a market for ideas?.” Industrial and Corporate Change 19.3 (2010): 805-837.

Glennon, Britta, et al. Does Employing Skilled Immigrants Enhance Competitive Performance? Evidence from European Football Clubs. No. w29446. National Bureau of Economic Research, 2021.

Hernandez, Exequiel, and Elena Kulchina. ”Immigrants and foreign firm performance.” Organization Science 31.4 (2020): 797-820.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. ”German Jewish émigrés and US invention.” American Economic Review 104.10 (2014): 3222-55.

Moser, Petra, and Shmuel San. ”Immigration, science, and invention. lessons from the quota acts.” Lessons from the Quota Acts (March 21, 2020) (2020).

Peri, Giovanni, and Chad Sparber. ”Task specialization, immigration, and wages.” American Economic Journal: Applied Economics 1.3 (2009): 135-69.

Roach, Michael, and Henry Sauermann. ”A taste for science? PhD scientists’ academic orientation and self-selection into research careers in industry.” Research policy 39.3 (2010): 422-434.

Roy, Andrew Donald. ”Some thoughts on the distribution of earnings.” Oxford economic papers 3.2 (1951): 135-146.

Sauermann, Henry, and Michael Roach. "Not all scientists pay to be scientists: PhDs' preferences for publishing in industrial employment." Research Policy 43.1 (2014): 32-47.

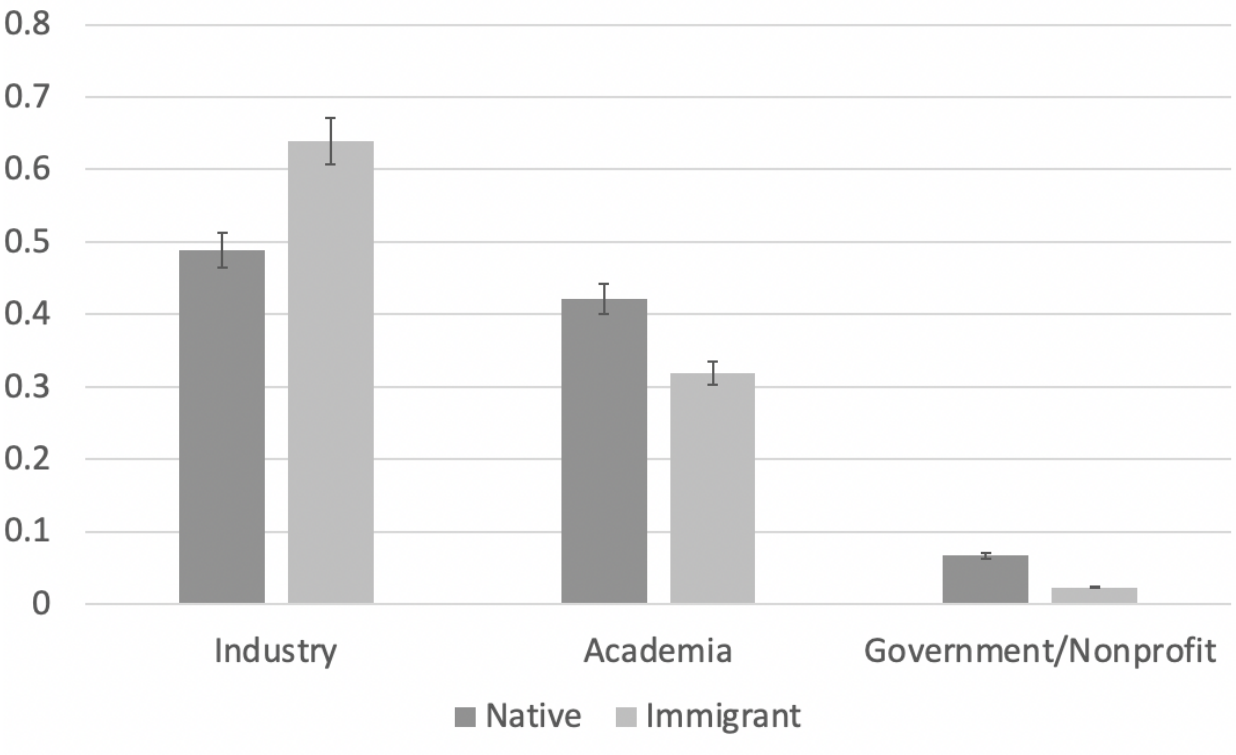
Stern, Scott. "Do scientists pay to be scientists?." Management science 50.6 (2004): 835-853.

Wang, Dan. "Activating cross-border brokerage: Interorganizational knowledge transfer through skilled return migration." Administrative Science Quarterly 60.1 (2015): 133-176.

Tables and Figures

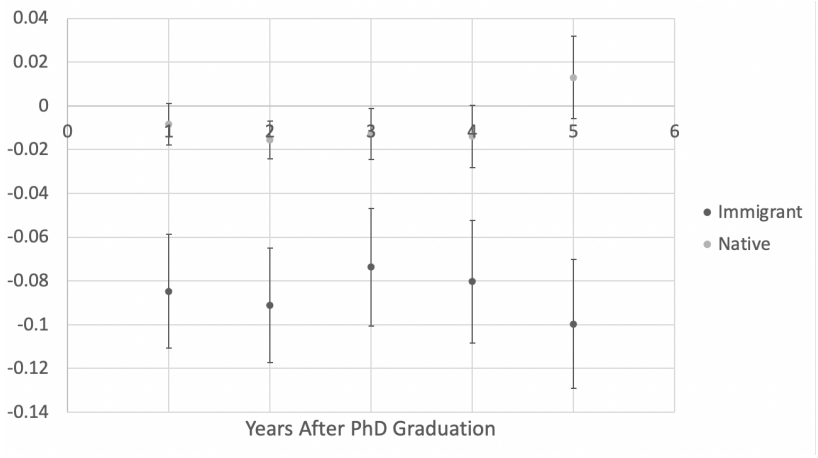
Figures

Figure 1: Job Type for Immigrant and Native Graduates

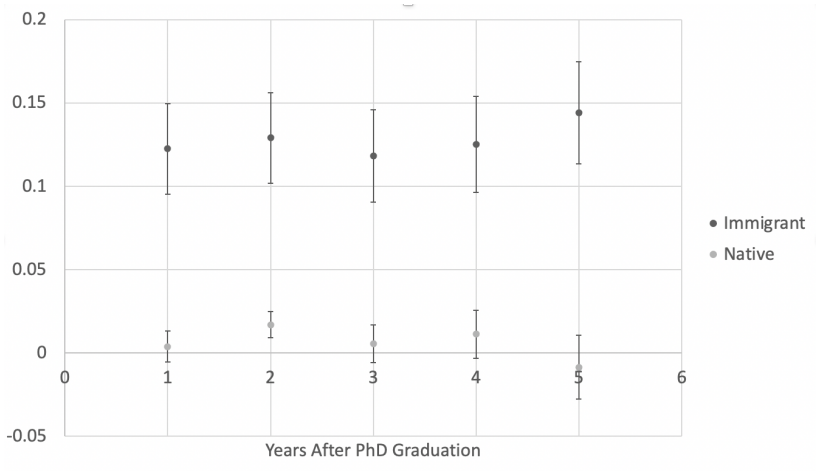


Notes: We plot the count of jobs, with 95% confidence intervals, started in industry, academia and government/non-profit amongst our sample PhD graduates.

Figure 2: Probability of Starting Academic or Industry Job Since PhD Graduation



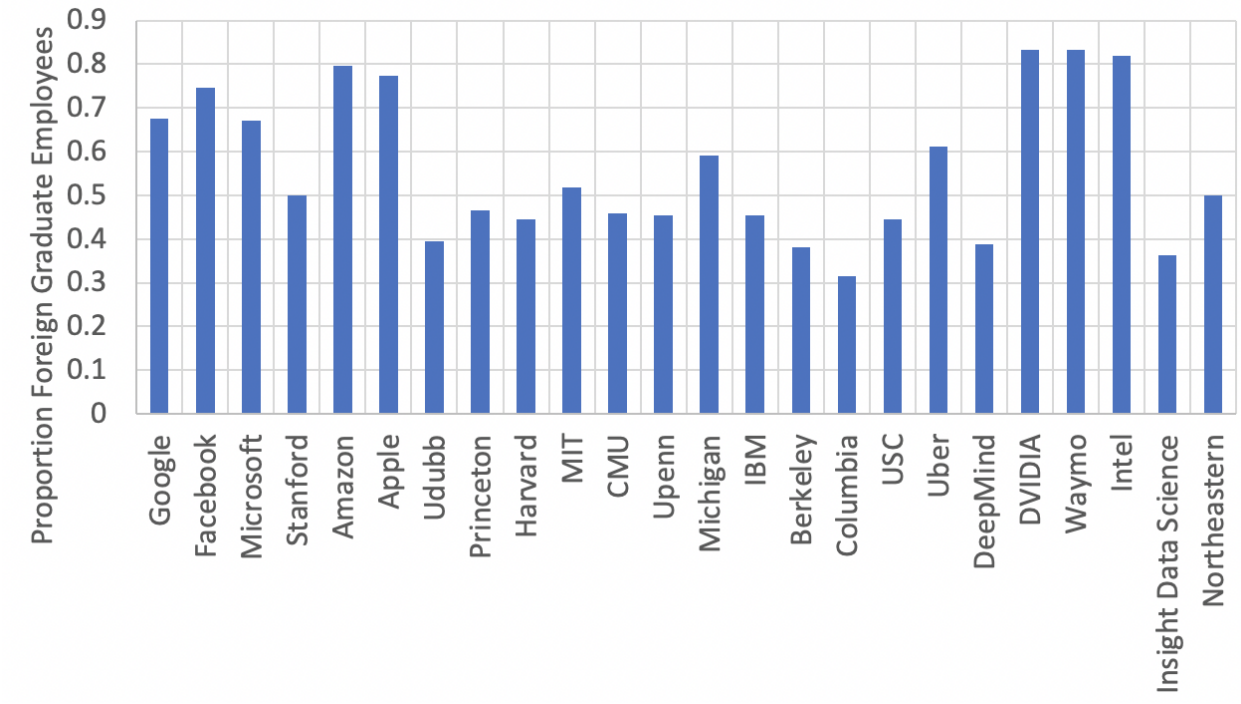
(a) Academic Job



(b) Industry Job

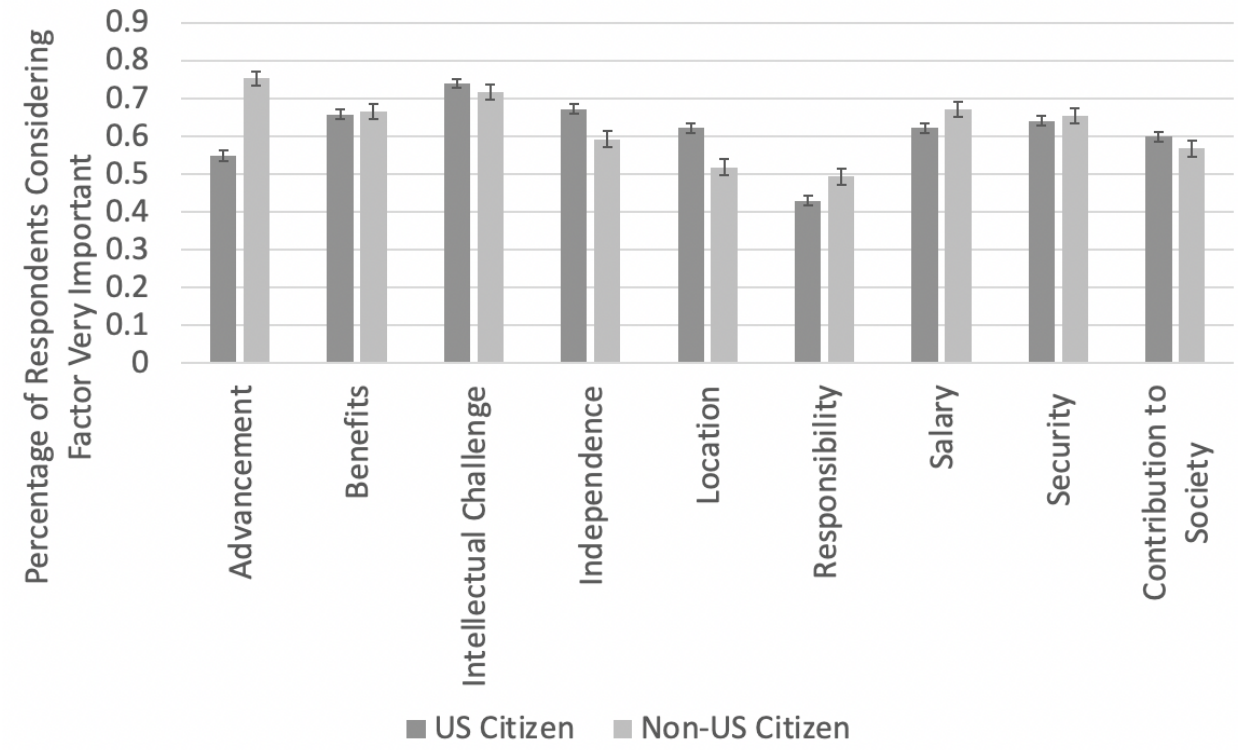
Notes: In panel a we plot the coefficients of the regression of the probability of having an academic job on immigrant interacted with each year after PhD graduation, controlling for department, institution, graduation year and publications before graduation. In panel b we plot the coefficients of the regression of the probability of having an industry job on immigrant interacted with each year after PhD graduation, controlling for department, institution, graduation year and publications before graduation. The omitted group is natives in year 0 since PhD graduation.

Figure 3: Proportion of Immigrants by Top Employers



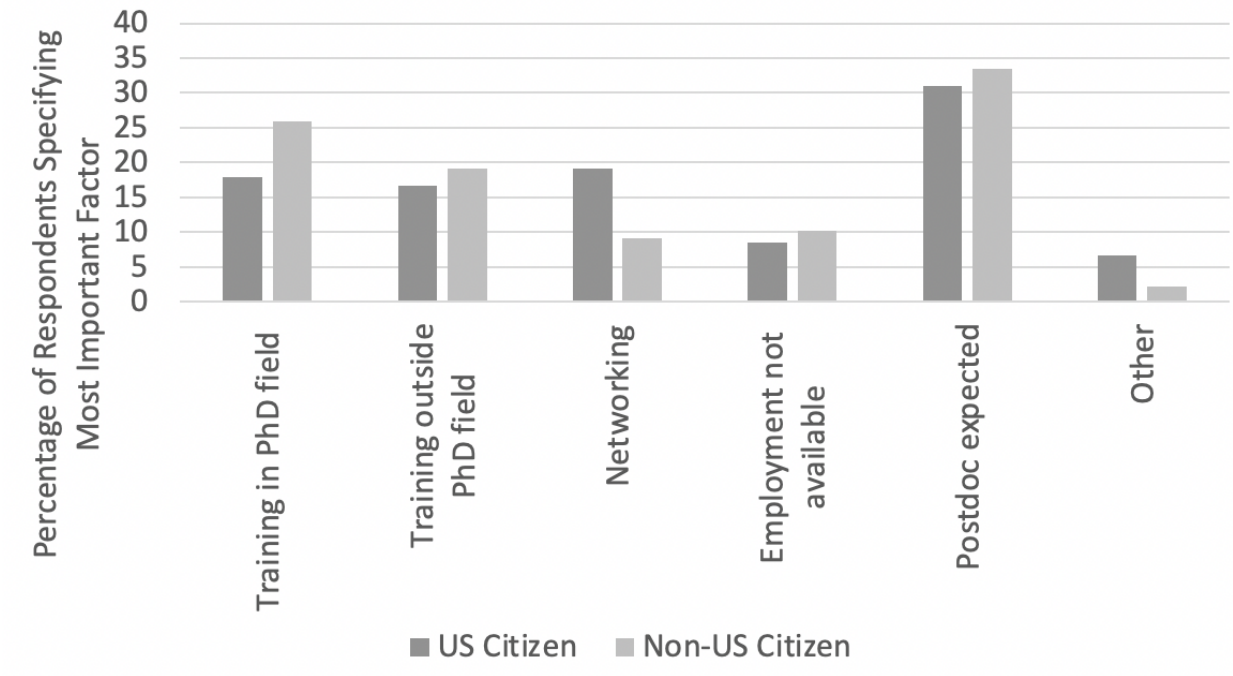
Notes: We plot the proportion of immigrant graduates in our sample employed by each organization.

Figure 4: Proportion of Respondents Considering Factor Very Important



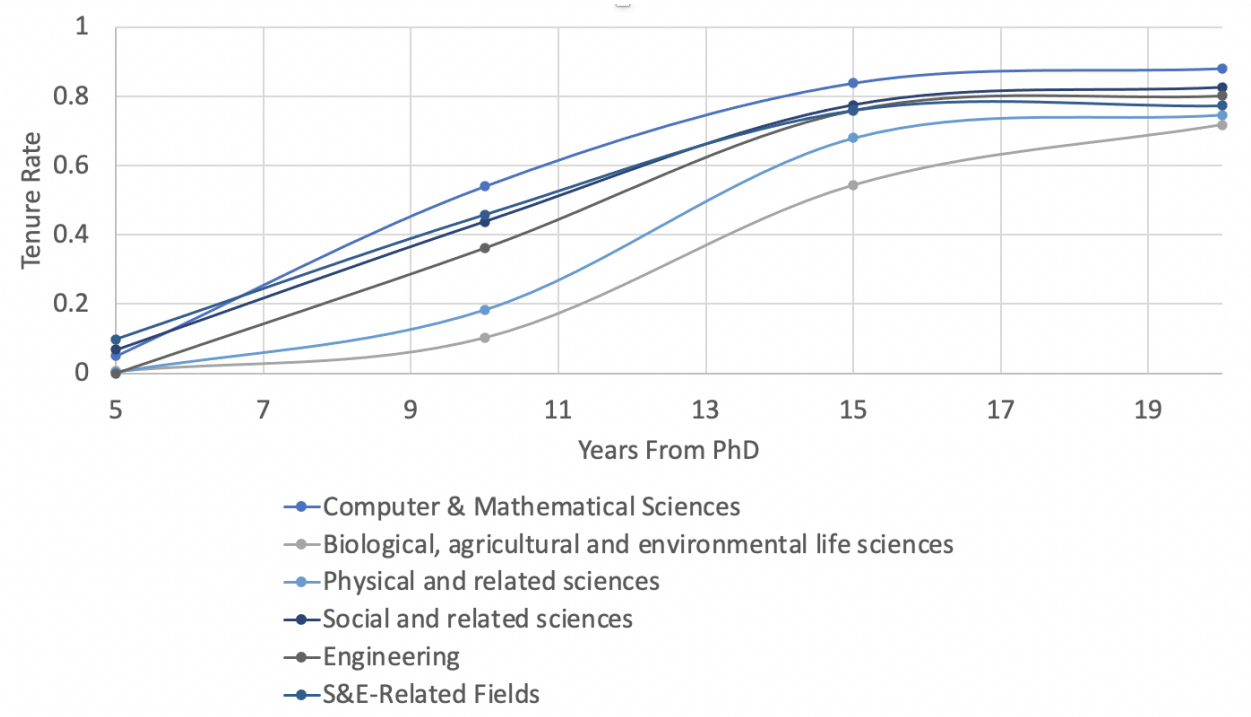
Notes: NSF data, 2015 graduates.

Figure 5: Proportion of Respondents Considering Factor Very Important in Doing PostDoc



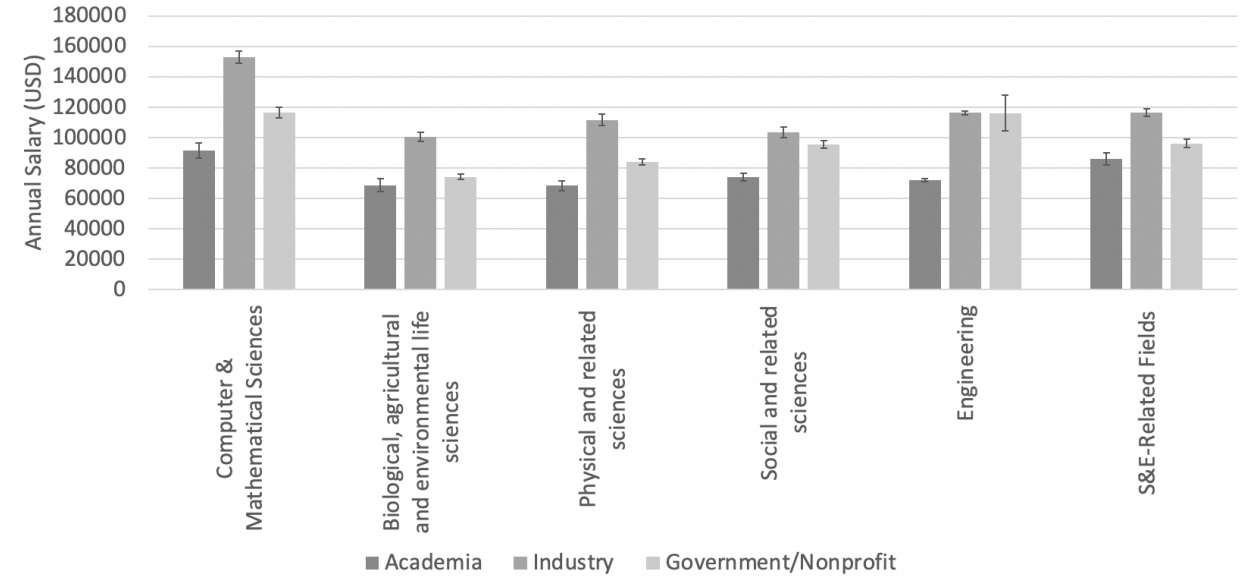
Notes: NSF data, 2015 graduates.

Figure 6: Tenure Rate by PhD Department



Notes: NSF data, postdocs and tenure track employees, proportion tenured within 5 year bracket post-PhD.

Figure 7: Salary by PhD Department



Notes: NSF data, 2015 graduates, 7,350 PhD graduates, average salary with 95% confidence intervals.

Table 1: Graduating Institution for Study Sample AI Graduates

	Nb Sample Scientists	Immigrant Percentage
Overall	1,769	0.53
Institution		
California Institute of Technology	14	0.71
Carnegie Mellon University	54	0.63
Columbia University	80	0.50
Cornell University	68	0.50
Georgia Institute of Technology	61	0.57
Harvard University	34	0.44
Massachusetts Institute of Technology	174	0.43
Princeton University	62	0.48
Stanford University	193	0.37
The University of Texas at Austin	9	0.67
The University of Wisconsin - Madison	69	0.58
University of California, Berkeley	154	0.42
University of California, Los Angeles	64	0.63
University of Illinois	97	0.69
University of Maryland, College Park	131	0.68
University of Massachusetts Amherst	129	0.53
University of Michigan	78	0.58
University of Pennsylvania	70	0.50
University of Southern California	96	0.81
University of Washington	132	0.42
Department		
Engineering	739	0.59
Mathematics and Computer Science	537	0.55
Life Sciences	119	0.29
Psychology and Social Sciences	85	0.28
Physical Sciences	75	0.31
Other Non-S&E	50	0.50
Humanities and Arts	41	0.44
Other Science and Engineering	19	0.47
Education	18	0.28

Notes: This sample consists of 1,769 artificial intelligence US PhD graduates.

Table 2: Descriptive Statistics

	Native (N = 837) (1)		Immigrant (N = 932) (2)		(2)-(1)
	mean	std. dev.	mean	std. dev.	mean (std. dev.)
Undergrad Graduation Year	2004	8.24	2006	6.78	1.18 (2.22)
PhD Graduation Year	2016	1.41	2016	1.38	0.024 (0.066)
Top 10 PhD Institution	0.66	0.47	0.50	0.50	-0.16*** (0.023)
Pre-PhD graduation number of publications	27.22	180.39	140.29	555.72	113.07*** (20.10)
Pre-PhD graduation number of patents	1.89	19.04	17.57	196.98	15.70** (6.83)
Pre-PhD graduation advisor patents	3.96	29.40	2.80	16.51	-1.15 (1.12)
Work in US post PhD graduation	0.95	0.21	0.92	0.27	-0.03 (0.040)

Notes: This sample consists of 1,769 artificial intelligence US PhD graduates.

Table 3: Occupational Choice for PhD Graduates

	Any Job				First Job	Second Job
	(1)	(2)	(3)	(4)	(5)	(6)
DV: Industry Employment						
Immigrant	0.1493***	0.1360***	0.1206***	0.1064***	0.1037***	0.1460***
	(0.024)	(0.024)	(0.026)	(0.026)	(0.029)	(0.048)
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5517	0.6000
DV: Academic Employment						
Immigrant	-0.1019***	-0.0882***	-0.0775***	-0.0634***	-0.0775***	-0.1063**
	(0.022)	(0.022)	(0.024)	(0.024)	(0.028)	(0.044)
Mean of Dep. Variable	0.3675	0.3675	0.3675	0.3675	0.3878	0.3318
Total Observations	2743	2743	2743	2743	1769	660
Graduation Year FE	X	X	X	X	X	X
PhD Institution FE		X	X	X	X	X
Pre-Graduation Publications			X	X	X	X
PhD Department FE				X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are dummy variables that take the value of 1 if the graduate takes an industry or academic job, 0 otherwise.

[b] Heteroskedastic robust standard errors, clustered at the level of the PhD advisor, are given in parentheses.

Table 4: Occupational Choice for PhD Graduates

	Any Industry Job				
	(1)	(2)	(3)	(4)	(5)
Non-European Name Ethnicity	0.1771***	0.1574***	0.1392***	0.1236***	
	(0.024)	(0.023)	(0.026)	(0.026)	
Non-European Name Ethnicity, US undergraduate					0.0600
					(0.037)
Non-European Name Ethnicity, non US undergraduate					0.1503***
					(0.029)
Total Observations	2743	2743	2743	2743	2743
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5680
Graduation Year FE	X	X	X	X	X
PhD Institution FE		X	X	X	X
Pre-Graduation Publications			X	X	X
PhD Department FE				X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are dummy variables that take the value of 1 if the graduate takes an industry job, 0 otherwise.

[b] Heteroskedastic robust standard errors, clustered at the level of the PhD advisor, are given in parentheses.

Table 5: Organization Type and Role

	Employer			Role			
	Multinational Corporation	Startup Company	Government Employee	Postdoc Fellowship	Teaching Role	Management–Industry	Research – Industry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Immigrant	2.1402***	1.01265	0.79749	0.9329	0.4852	0.8955	1.5751**
	(0.226)	(0.134)	(0.290)	(0.184)	(0.161)	(0.259)	(0.280)
Total Observations	2743	2743	2743	2743	2743	2743	2743
Mean of Dep. Variable	0.3751	0.1619	0.2741	0.1280	0.0252	0.0379	0.3146
Graduation Year FE	X	X	X	X	X	X	X
PhD Institution FE	X	X	X	X	X	X	X
Pre-Graduation Publications	X	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects multinomial logistic regression specifications. The comparison occupational choice in columns (1-3) is working in a university, and in columns (4-7) the comparison role is tenure track university role.
[b] Heteroskedastic robust standard errors are given in parentheses.

Table 6: Occupation Choice for PhD Graduates by Country of Origin

	Any Industry Job			
	(1)	(2)	(3)	(4)
Immigrant	0.0684**	0.1048***	0.1246***	0.0654**
	(0.029)	(0.027)	(0.027)	(0.033)
Chinese Immigrant	0.1513***			0.1571***
	(0.036)			(0.040)
Indian Immigrant		0.0097		0.0548
		(0.044)		(0.047)
High-Income Country Immigrant			-0.0968*	-0.0393
			(0.053)	(0.056)
Total Observations	2743	2743	2743	2743
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680
Graduation Year FE	X	X	X	X
PhD Institution FE	X	X	X	X
Pre-Graduation Publications	X	X	X	X
PhD Department FE	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are dummy variables that take the value of 1 if the graduate takes an industry job, 0 otherwise.

[b] Heteroskedastic robust standard errors, clustered at the level of the PhD advisor, are given in parentheses.

Table 7: Heterogeneity in Occupation Choice for PhD Graduates

	Any Industry Job							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigrant	0.1064***	0.1063***	0.1027***	0.0928**	0.1304***	0.1016***	0.1374***	26.9717
	(0.026)	(0.026)	(0.027)	(0.045)	(0.038)	(0.027)	(0.042)	(34.558)
Advisor Nb. Patents		-0.0005	-0.0007***					
		(0.000)	(0.000)					
Immigrant × Advisor Nb. Patents			0.0010					
			(0.001)					
Immigrant × Institution Foreign Ratio				0.0110				
				(0.030)				
Immigrant × Pre-Graduation Pubs					-0.0132			
					(0.014)			
Immigrant × Pre-Graduation Patents						0.0177		
						(0.015)		
Immigrant × Top 10 University							-0.0522	
							(0.053)	
Immigrant × Graduation Year								-0.0133
								(0.017)
Total Observations	2743	2743	2743	2743	2743	2743	2743	2743
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5680	0.5680	0.5680	0.5680
Graduation Year FE	X	X	X	X	X	X	X	X
PhD Institution FE	X	X	X	X	X	X	X	X
Pre-Graduation Publications	X	X	X	X	X	X	X	X
PhD Department FE	X	X	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are dummy variables that take the value of 1 if the graduate takes an academic job, 0 otherwise.

[b] Heteroskedastic robust standard errors, clustered at the level of the PhD advisor, are given in parentheses.

Table 8: Heterogeneity in Occupation Choice for PhD Graduates

	Any Industry Job					
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant	0.1064***	0.0972***	0.1159***	0.0816***	0.1062***	0.0206
	(0.026)	(0.033)	(0.027)	(0.031)	(0.027)	(0.059)
Immigrant × Engineering PhD		0.0218				0.0985
		(0.051)				(0.071)
Immigrant × Life Sciences PhD			-0.1576*			-0.0617
			(0.094)			(0.109)
Immigrant × Computer Science PhD				0.0795		0.1403*
				(0.054)		(0.074)
Immigrant × Physical Science PhD					0.0029	0.0890
					(0.118)	(0.127)
Total Observations	2743	2743	2743	2743	2743	2743
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5680	0.5680
Graduation Year FE	X	X	X	X	X	X
PhD Institution FE	X	X	X	X	X	X
Pre-Graduation Publications	X	X	X	X	X	X
PhD Department FE	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are dummy variables that take the value of 1 if the graduate takes an academic job, 0 otherwise.
[b] Heteroskedastic robust standard errors, clustered at the level of the PhD advisor, are given in parentheses.

Table 9: Publication and Patenting Output for PhD Graduates

	Annual Number of Publications				Annual Number of Patents				SNIP Weighted Number of Publications	Number of Top Publications
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Immigrant	0.0299		0.0702**	0.1158***	0.0115		0.0052	-0.0275*	0.0065	-0.0002
	(0.033)		(0.030)	(0.044)	(0.017)		(0.016)	(0.016)	(0.028)	(0.004)
Industry Job		-0.3716***	-0.3785***	-0.3384***		0.0609***	0.0604***	0.0317*	-0.0251	0.0001
		(0.030)	(0.030)	(0.037)		(0.014)	(0.014)	(0.017)	(0.020)	(0.003)
Immigrant × Industry Job				-0.0839				0.0600**	0.0047	0.0046
				(0.056)				(0.030)	(0.036)	(0.005)
Total Observations	9745	9745	9745	9745	9745	9745	9745	9745	9745	9745
Mean of Dep. Variable	24.1336	24.1336	24.1336	24.1336	2.2842	2.2842	2.2842	2.2842	4.3474	0.0563
Graduation Year FE	X	X	X	X	X	X	X	X	X	X
PhD Institution FE	X	X	X	X	X	X	X	X	X	X
Pre-Graduation Publications	X	X	X	X	X	X	X	X	X	X
Pre-Graduation Patents	X	X	X	X	X	X	X	X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are inverse hyperbolic sine transformed publication and patent counts. SNIP weighted publications are publications weighted by the journal's source normalized impact per paper, a measure of a journal's reach. Top publications are those in journals in the top 95th percentile of sample SNIP.

[b] Heteroskedastic robust standard errors, clustered at the level of the PhD advisor, are given in parentheses.

Appendix

Table A1: Occupational Choice for PhD Graduates - standard errors clustered at the institutional level

	Any Industry Job				First Job Industry	Second Job Industry
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant	0.1493***	0.1360***	0.1206***	0.1064***	0.1106***	0.1476**
	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)	(0.055)
Total Observations	2743	2743	2743	2743	1769	660
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5495	0.6000
Graduation Year FE	X	X	X	X	X	X
PhD Institution FE		X	X	X	X	X
Pre-Graduation Publications			X	X	X	X
PhD Department FE				X	X	X

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: [a] Estimates stem from fixed effects ordinary least squares specifications in which dependent variables are dummy variables that take the value of 1 if the graduate takes an industry, 0 otherwise.

[b] Heteroskedastic robust standard errors, clustered at the level of the PhD institution, are given in parentheses.

Table A2: Missing Data

	Nb Sample Scientists	Nb Scientists Incl Missing Scientists
Overall	1,769	3,123
Average graduation year	2016	2016
Proportion non-European ethnic name	0.60	0.59
Institution		
California Institute of Technology	14	16
Carnegie Mellon University	54	82
Columbia University	80	159
Cornell University	68	120
Georgia Institute of Technology	61	89
Harvard University	34	64
Massachusetts Institute of Technology	174	287
Princeton University	62	103
Stanford University	193	341
The University of Texas at Austin	9	12
The University of Wisconsin - Madison	69	143
University of California, Berkeley	154	256
University of California, Los Angeles	64	205
University of Illinois	97	176
University of Maryland, College Park	131	226
University of Massachusetts Amherst	129	251
University of Michigan	78	126
University of Pennsylvania	70	110
University of Southern California	96	146
University of Washington	132	211

Notes: This sample consists of 1,769 artificial intelligence US PhD graduates out of the 3,123 scientists who graduated from top AI PhD programs for whom we could obtain information on undergraduate institution and post-PhD graduation employment history.