

Immigrant scientist career choices and the shape of science: Evidence from Artificial Intelligence PhDs

Caroline Fry, University of Hawai'i

Britta Glennon, Wharton and NBER

TECHNOLOGY EXECUTIVE COUNCIL

Why some college professors are adopting ChatGPT AI as quickly as students

PUBLISHED SUN, APR 2 2023-9:13 AM EDT | UPDATED MON, APR 3 2023-9:58 AM EDT



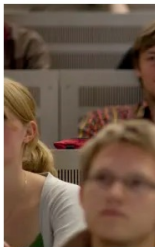
Carolyn Chun

SHARE [f](#) [t](#) [in](#) [✉](#)

Advertisement

KEY POINTS

- A recent study found that many professors are using ChatGPT to grade essays.
- One of the main reasons is that ChatGPT can generate high-quality text that is difficult to detect as AI-generated.
- But use of ChatGPT by students has also increased, leading to concerns about academic integrity.



HOME > TECH

Two professors who say they caught students cheating on essays with ChatGPT explain why AI plagiarism can be hard to prove

Beatrice Nolan Jan 13, 2023, 10:00 PM



Photo by Getty Images

How to help their students

INSIDER
Newsletters

Personal Finance Newsletter
Learn how to manage your money

Articles delivered to your inbox biweekly

[SIGN UP FOR FREE](#)

TECHNOLOGY

The College Essay Is Dead

Nobody is prepared for how AI will transform academia.

By Stephen Marche

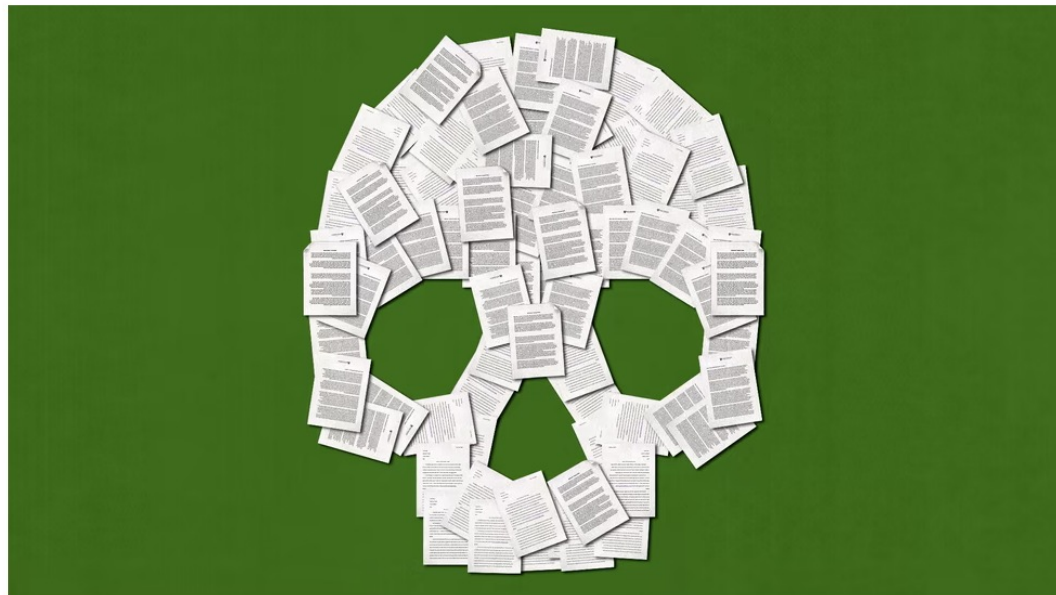
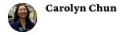


Photo by Paul Spiller / The Atlantic; Getty Images

TECHNOLOGY EXECUTIVE COUNCIL

Why some college professors are adopting ChatGPT AI as quickly as students

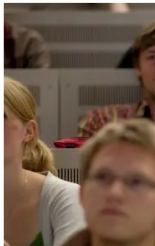
PUBLISHED SUN, APR 2 2023



Carolyn Chun

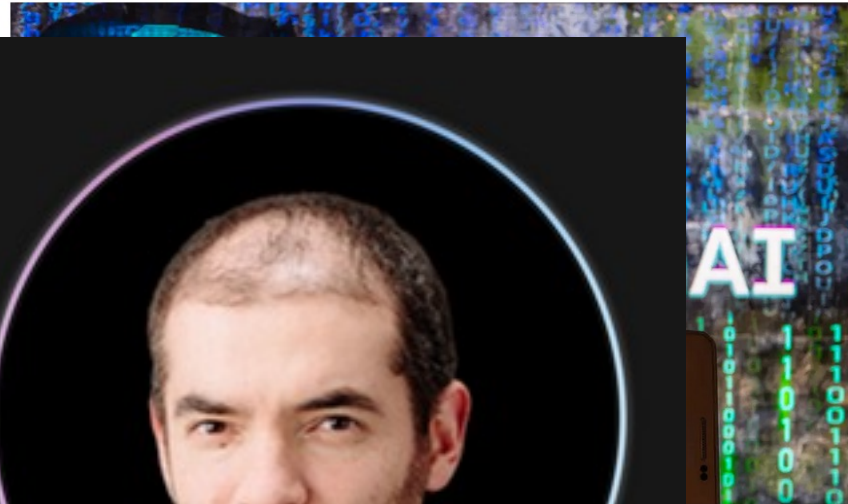
KEY POINTS

- A recent study found that many professors are in the dark about AI tools like ChatGPT.
- One of the reasons is that many professors are not tech-savvy.
- But use of AI tools can be critical to the success of a course.



Two professors who say they caught students cheating on essays with ChatGPT explain why AI plagiarism can be hard to prove

Beatrice Nolan Jan 13, 2023, 10:00 PM



Ilya Sutskever

Co-Founder and Chief Scientist of OpenAI

INSIDER Newsletters

Personal Finance Newsletter

Learn how to manage your money

Articles delivered to your inbox biweekly

[SIGN UP FOR FREE](#)



Overview of paper

Research Questions:

- Are immigrants more likely than natives to choose industry over academia?
- What are the implications of this career choice for the production and diffusion of knowledge?

Overview of paper

Research Questions:

- Are immigrants more likely than natives to choose industry over academia?
- What are the implications of this career choice for the production and diffusion of knowledge?

Setting & Data:

- Sample of Artificial Intelligence US PhDs
 - career trajectory information (CSET)
 - publication data (OpenAlex)

Overview of paper




Research Questions:

- Are immigrants more likely than natives to choose industry over academia?
- What are the implications of this career choice for the production and diffusion of knowledge?

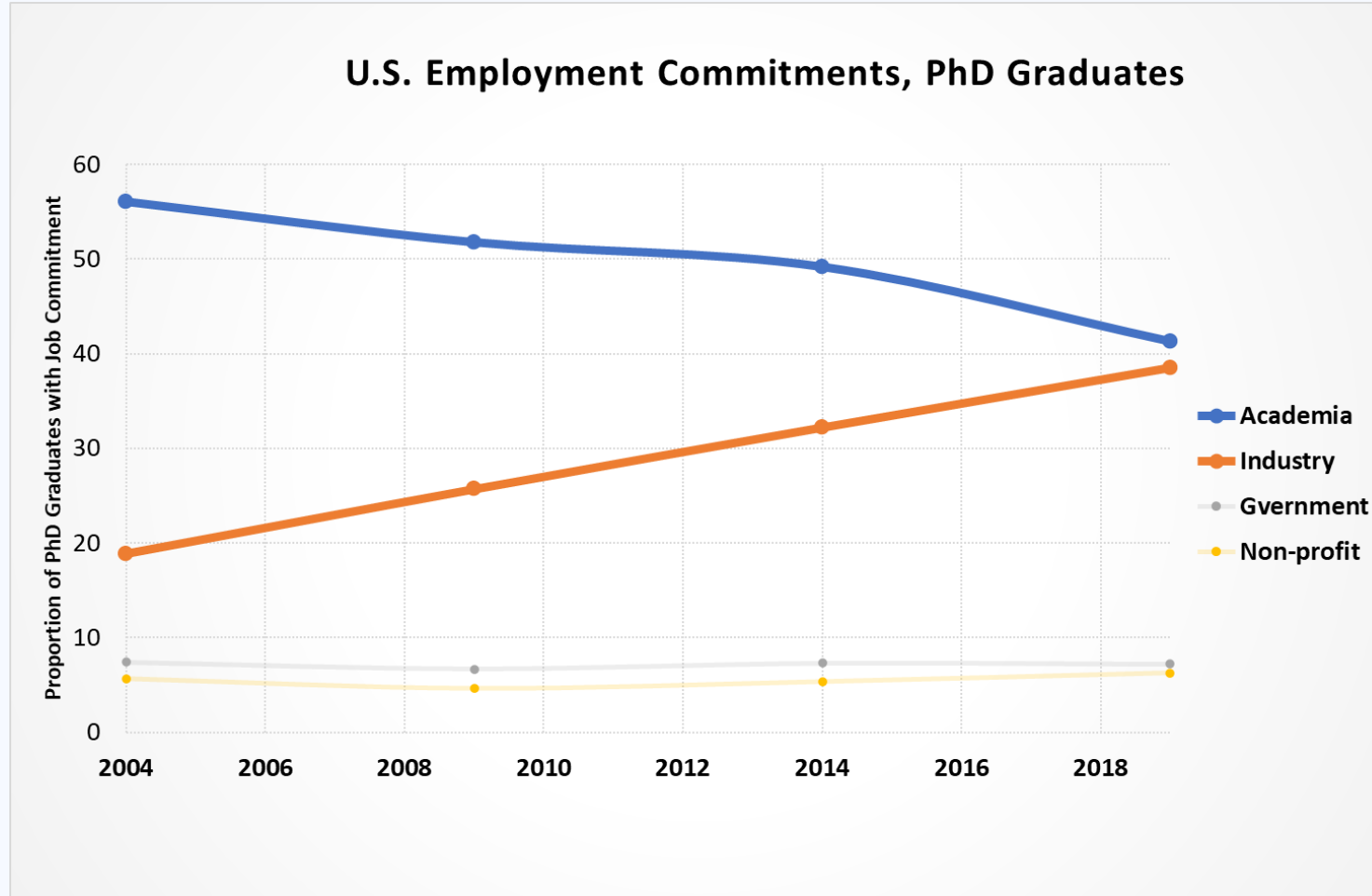
Setting & Data:

- Sample of Artificial Intelligence US PhDs
 - career trajectory information (CSET)
 - publication data (OpenAlex)

Main Findings:

- Immigrant PhDs in industry 
- Rate and range of knowledge production 
- Global diffusion of knowledge 

PhD graduates across disciplines are increasingly likely to go into industry rather than academia



Data: National Center for Science and Engineering Statistics, Survey of Earned Doctorates.

But PhD graduates are not a homogenous body

Immigrants make up a large, growing fraction of S&E PhD students in the US

- 39.4% overall and 51% in computer science in 2021 (NSF)
- Up from 28.1% in 2011

But PhD graduates are not a homogenous body

Immigrants make up a large, growing fraction of S&E PhD students in the US

- 39.4% overall and 51% in computer science in 2021 (NSF)
- Up from 28.1% in 2011

And immigrants might have different preferences and constraints than non-immigrants...

The academia-industry choice and a “taste for science”: What we know

What affects the industry-academia tradeoff?

Taste for science/preference for publishing (*Agarwal and Ohyama 2012; Stern 2004; Roach and Sauermann 2010, 2014*)

Preferences for money (*Agarwal and Ohyama 2012, Roach and Sauermann 2010*)

Industry/University demand for basic versus applied research (*Agarwal and Ohyama 2012*)

Sorting by ability (*Agarwal and Ohyama 2012*)

Complements available (*Agarwal and Ohyama 2012*)

The academia-industry choice and a “taste for science”: What we know

What affects the industry-academia tradeoff?

Taste for science/preference for publishing (*Agarwal and Ohyama 2012; Stern 2004; Roach and Sauermann 2010, 2014*)

Preferences for money (*Agarwal and Ohyama 2012, Roach and Sauermann 2010*)

Industry/University demand for basic versus applied research (*Agarwal and Ohyama 2012*)

Sorting by ability (*Agarwal and Ohyama 2012*)

Complements available (*Agarwal and Ohyama 2012*)

But this literature has not distinguished how the “taste for science” might differ for immigrants and natives...

Immigrant graduate students prioritize different factors in their job search

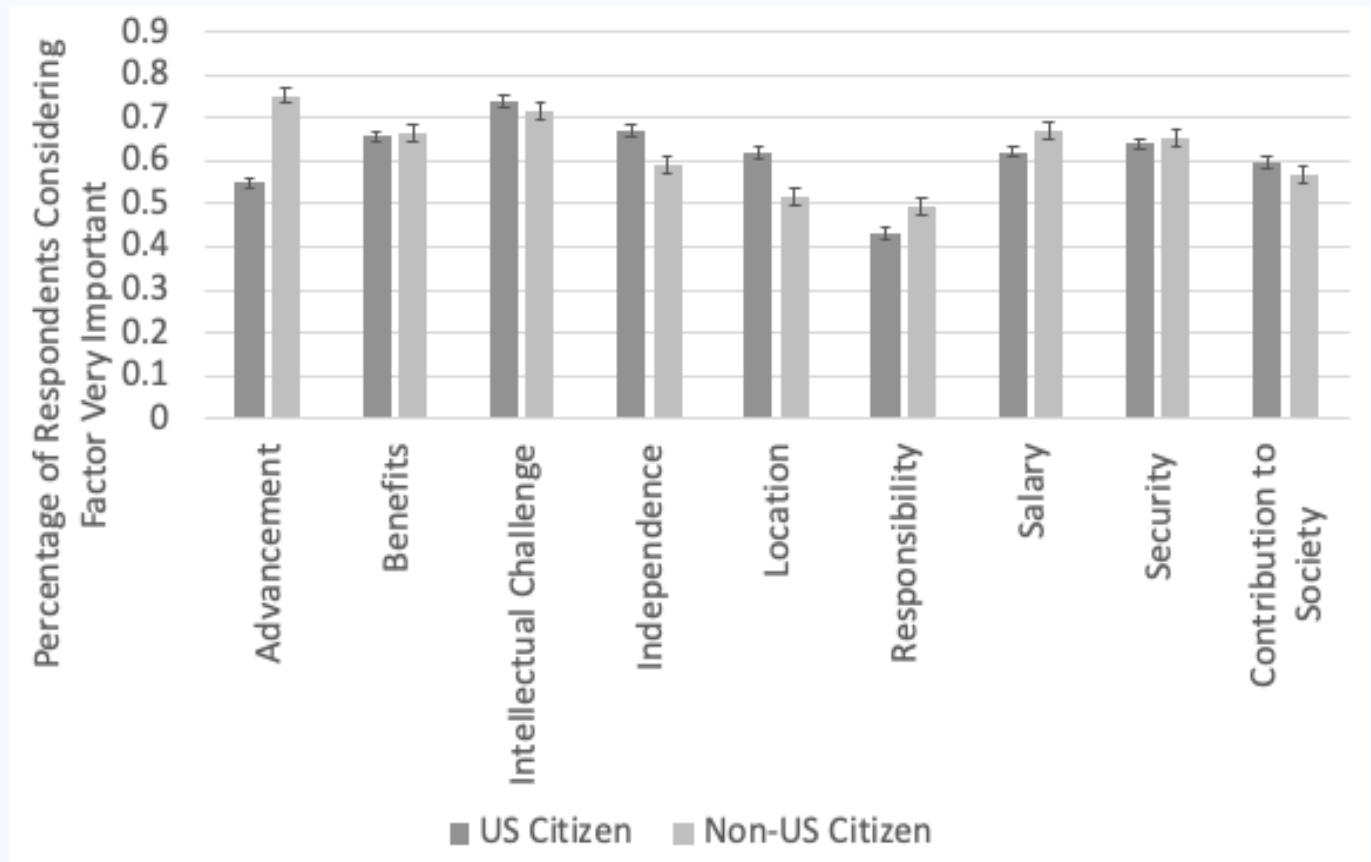
According to NSF...

Immigrants were more likely to place value on:

- Career advancement possibilities
- Responsibility
- Salary

Immigrants were less likely to place value on:

- Independence
- Job location



Source; 2015 NSF SED data, raw data plotted with 95% confidence intervals

And immigrants face different institutional constraints

For-profit companies face an H-1B/EB-1 visa cap, while non-profit universities do not.

Some companies (especially start-ups) are unwilling/unable to sponsor H-1B/EB-1 visas due to high costs and uncertainty

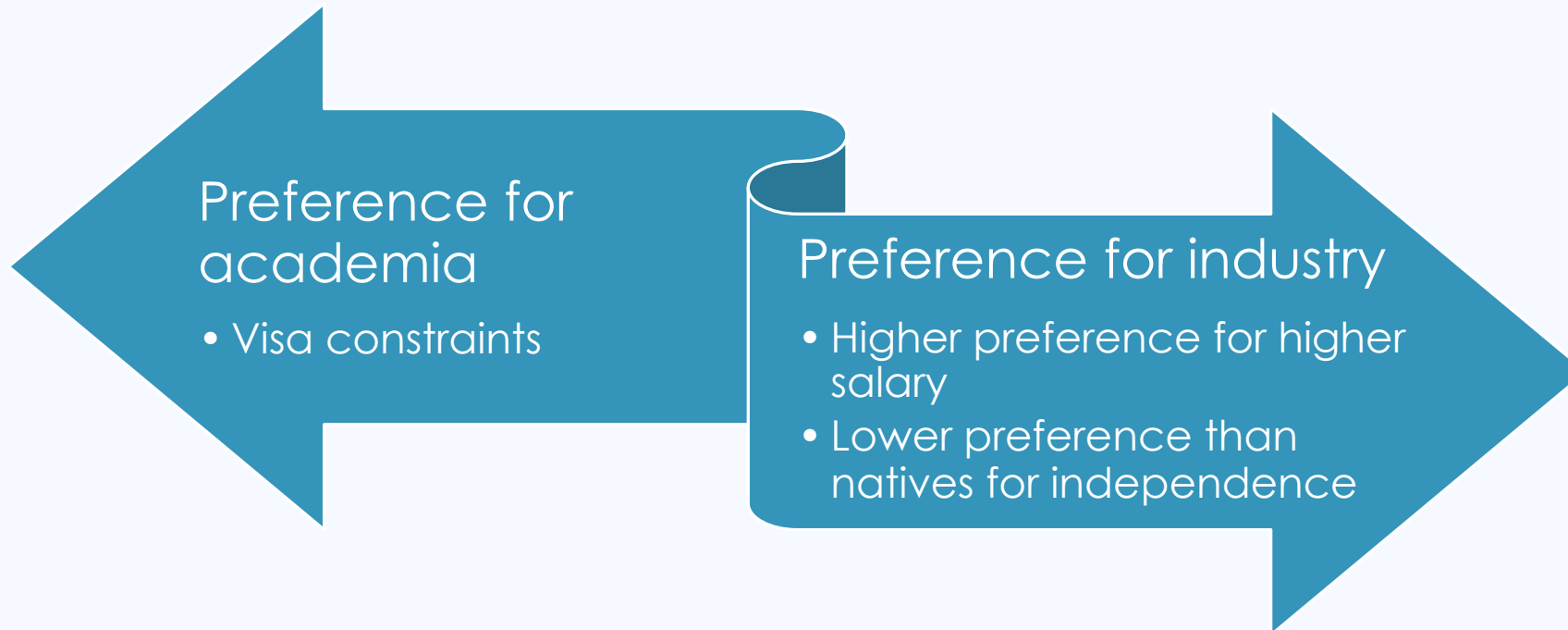
H-1B petitions are distributed by lottery; fewer than half are granted in any given year

→ expect visa constraints to push immigrants into academia as opposed to industry

But the literature has found mixed evidence on the effect of immigration constraints on occupational choice

Finding	Choice being made	Comparison group	Sample for choice	Papers
Constraints → lower likelihood of working for or forming a startup	Within industry	Natives vs Immigrants (Roach and Skrentny 2019; Agarwal et al 2023) Chinese and Indian vs other nationalities (Diethorn 2022)	PhD graduates (Roach and Skrentny 2019; Diethorn 2023) All graduates (Agarwal et al 2022)	<ul style="list-style-type: none"> • Roach and Skrentny 2019 • Diethorn 2023 • Agarwal et al 2023
Constraints → higher likelihood of academia	Industry vs academia	International students pre- and post-2004	All foreign-born college graduates	<ul style="list-style-type: none"> • Amuedo-Dorantes and Furtado (2019)
no strong effect of visa constraints on occupational choice within the US	Industry vs academia	Chinese and Indian vs other nationalities	International STEM PhD graduates who stay in the US	<ul style="list-style-type: none"> • Kahn and MacGarvie (2020)

It is therefore ex-ante unclear whether immigrants would be more or less likely to go into industry



RQ1: Are immigrants more likely than natives to choose industry over academia?

Why should we care about where immigrant PhD graduates work?

Possible impact on the *rate* of science

- Proportion of immigrants in STEM doctorate programs has grown over time (61% in 2021 - NSF)
- Immigrants have a disproportionate impact on innovation (Bernstein et al 2022; Gaule and Piacentini (2013); Hunt and Gauthier-Loiselle 2010)
- Industry tends to focus on private production of knowledge

Possible impact on the *direction* of science

- Demographic characteristics shape the topics inventors work on (e.g. Koning Samila and Ferguson 2020, 2021; Nielsen et al 2017)
- Migrants have access to unique knowledge from their home country (e.g. Choudhury and Kim 2019; Agrawal et al 2011; Moser Voena Waldinger 2014)
- Industry tends to focus on applied, profit-oriented research

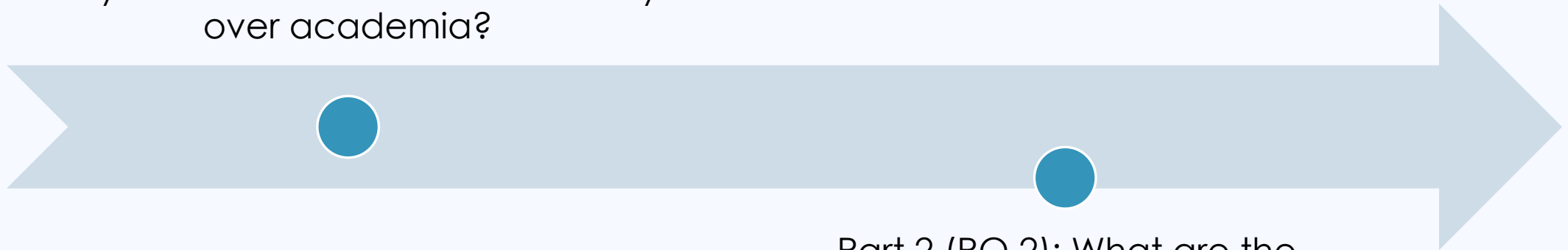
Possible impact on the *diffusion* of science

- Immigrants serve as a channel for global knowledge transfer (Agrawal et al 2011; Agrawal et al 2008; Bahar 2020; Ganguli 2015; Kerr 2008; Saxenian 2005; Kahn and MacGarvie 2012)
- Inventors pay different attention to discoveries made in academia vs industry (Bikard 2018; Bikard and Mrax 2019)

RQ2: What are the implications of immigrant PhD graduate career choice for the production and diffusion of knowledge?

Roadmap

Part 1 (RQ 1): Are immigrants more likely than natives to choose industry over academia?



Part 2 (RQ 2): What are the implications of immigrant career choice for the production & dissemination of science

Setting & data

Sample and data

Sample:

- Center for Security and Emerging Technology (CSET) database of the career history of 1,769 graduates of US PhD programs whose dissertations pertained to Artificial Intelligence (AI)
 - 20 highest ranking US programs for AI-related fields
 - Graduates between 2014 & 2018
 - Longitudinal data, employment record for up to 6 years after completion of PhD
- Defining immigrants:
 - Did they complete their undergraduate degree outside of the US? (preferred definition)
 - ~50% of our sample (20% are Chinese, 8% are Indian)
 - Name ethnicity (robustness)

Merged with:

- Publication data (OpenAlex)
- Patent data (Patentsview)

Why AI?

1. Especially high number of foreign-born students (~ 51% in computer science PhD in 2021, NSF)
2. Important field with exceptional potential for altering the innovative landscape, broader economy, and society at large
3. Demand for AI talent has grown more quickly than supply (Ahmed Wahed and Thompson 2023), removing the demand-side factors and helping us to isolate the supply-side
4. Industry is especially influential in AI (Ahmed Wahed and Thompson 2023)

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	4.32	7.60	7.19	9.62	2.87*** (0.42)
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)

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

95% of the sample remain in the US post-PhD

	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	4.32	7.60	7.19	9.62	2.87*** (0.42)
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)

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

Immigrants had more publications and patents before graduation

	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	4.32	7.60	7.19	9.62	2.87*** (0.42)
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)

Table 1: Graduating Institution for Study Sample AI Graduates

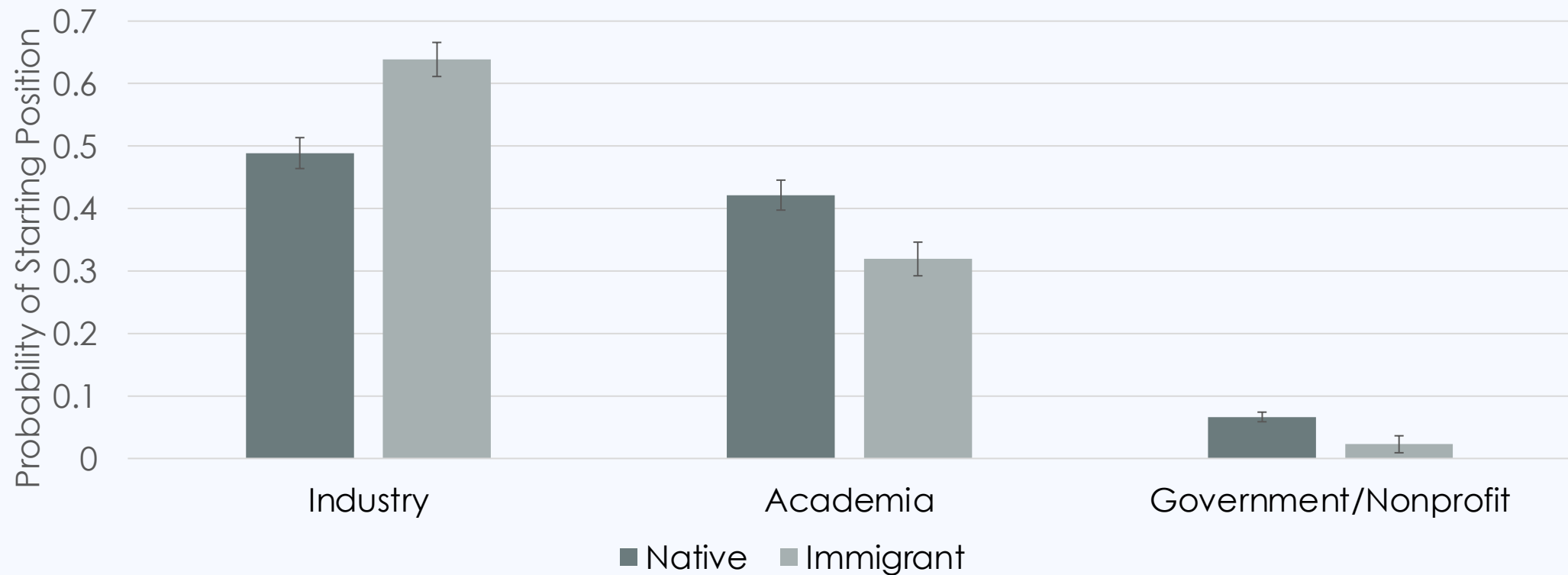
	Nb Sample Scientists	Immigrant Percentage
Overall	1,769	0.53
Institution		
California Institute of Technology	14	0.71
Carnegie Mellon University	54	0.63
Columbia University	80	0.50
Cornell University	68	0.50
Georgia Institute of Technology	61	0.57
Harvard University	34	0.44
Massachusetts Institute of 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

Roadmap

Part 1 (RQ 1): Are immigrants more likely than natives to choose industry over academia?

Part 2 (RQ 2): What are the implications of immigrant career choice for the production & dissemination of science

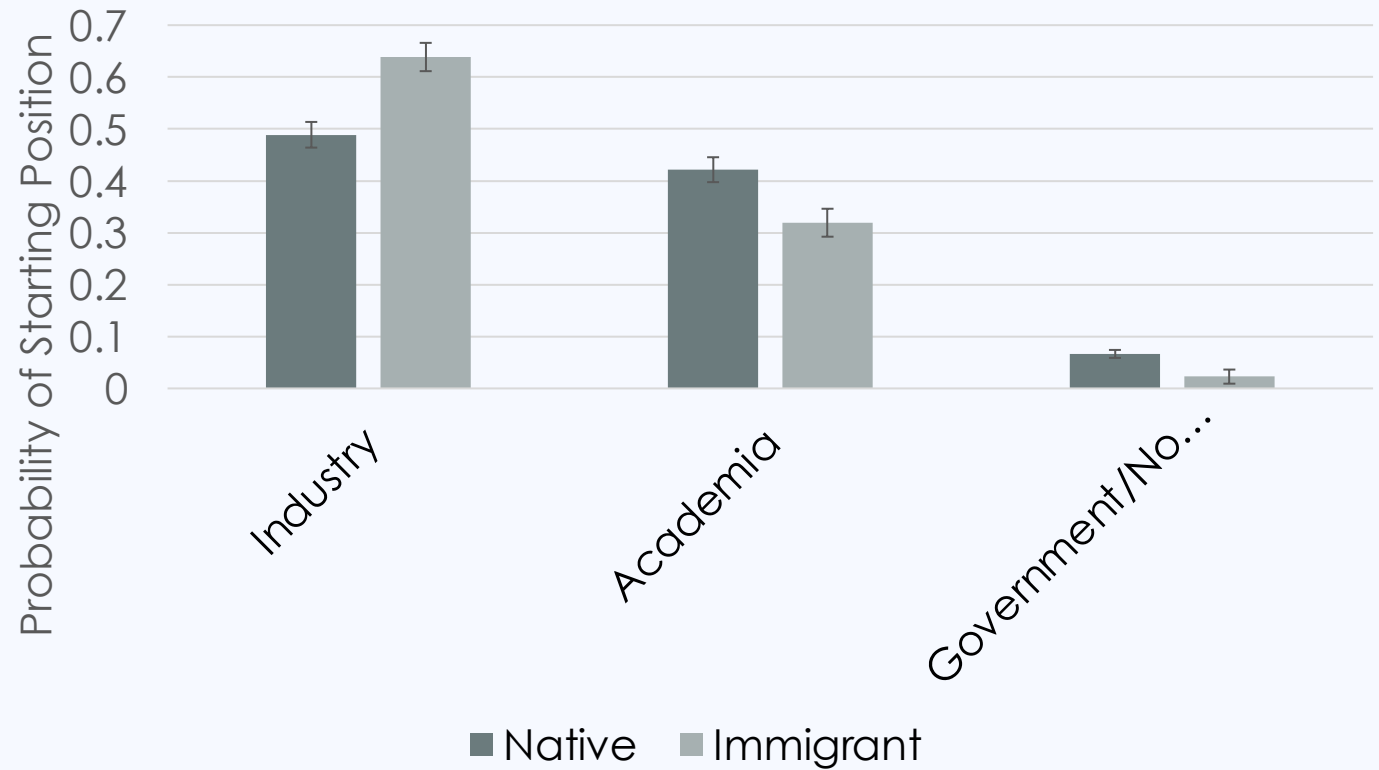
The raw data suggest that immigrant PhD graduates are more likely to choose industry



Notes: We plot raw career choice of sample PhD graduates, with 95% confidence intervals

The raw data suggest that immigrant PhD graduates are more likely to choose industry

But the results could be driven by selection into different departments or universities, for instance...



Econometric framework

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

- Dataset – person/job level
- Y_{ij} : Dummy variable 1 (industry job) 0 (not industry job)
- Controls for PhD publications, graduation year, institution, department
- SE's clustered at the advisor's level

Regression results also show that immigrant PhD graduates are more likely than natives to go into industry

	Any Industry Job				First Job Industry	Second Job Industry
	(1)	(2)	(3)	(4)	(5)	(6)
DV: Industry Employment						
Immigrant	0.1493*** (0.024)	0.1360*** (0.024)	0.1307*** (0.024)	0.1176*** (0.024)	0.1231*** (0.026)	0.1345*** (0.041)
Total Observations	2743	2743	2743	2743	1769	660
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5534	0.6015
DV: Academic Employment						
Immigrant	-0.1019*** (0.022)	-0.0882*** (0.022)	-0.0868*** (0.022)	-0.0730*** (0.022)	-0.0750*** (0.024)	-0.1043*** (0.039)
Total Observations	2743	2743	2743	2743	1769	660
Mean of Dep. Variable	0.3675	0.3675	0.3675	0.3675	0.3855	0.3348
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.

Regression results also show that immigrant PhD graduates are more likely than natives to go into industry

	Any Industry Job				First Job Industry	Second Job Industry
	(1)	(2)	(3)	(4)	(5)	(6)
DV: Industry Employment						
Immigrant	0.1493*** (0.024)	0.1360*** (0.024)	0.1307*** (0.024)	0.1176*** (0.024)	0.1231*** (0.026)	0.1345*** (0.041)
Total Observations	2743	2743	2743	2743	1769	660
Mean of Dep. Variable	0.5680	0.5680	0.5680	0.5680	0.5534	0.6015
DV: Academic Employment						
Immigrant	-0.1019*** (0.022)	-0.0882*** (0.022)	-0.0868*** (0.022)	-0.0730*** (0.022)	-0.0750*** (0.024)	-0.1043*** (0.039)
Total Observations	2743	2743	2743	2743	1769	660
Mean of Dep. Variable	0.3675	0.3675	0.3675	0.3675	0.3855	0.3348
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.

More specifically, immigrant graduates are more likely to go into a research role at a multinational corporation

	Employer			Role			
	Multinational Corporation	Startup Company	Government Employee	Postdoc Fellowship	Teaching Role	Management-Industry	Research – Industry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Immigrant	2.1585*** (0.209)	1.07193 (0.130)	0.9005 (0.300)	0.8514 (0.155)	0.38574 (0.123)	1.00027 (0.252)	1.5352** (0.247)
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

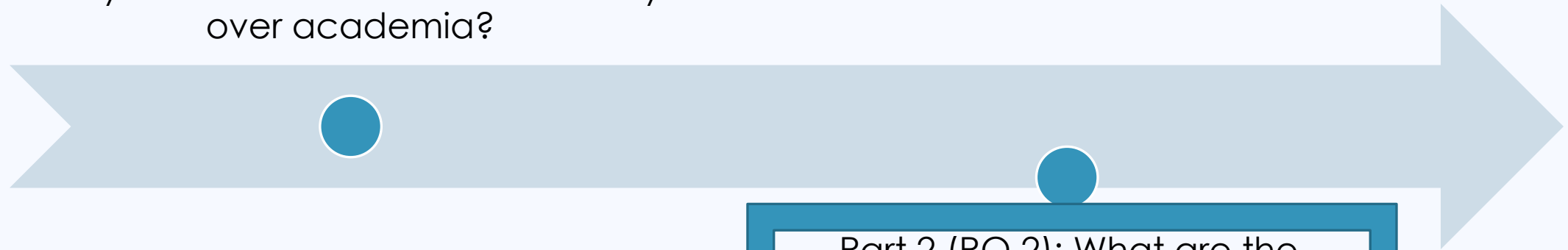
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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.

Roadmap

Part 1 (RQ 1): Are immigrants more likely than natives to choose industry over academia?



Part 2 (RQ 2): What are the implications of immigrant career choice for the production & dissemination of science

Implications of Career Choice

- Production of knowledge
 - Rate
 - Direction
- Diffusion of knowledge

Implications of Career Choice

- **Production of knowledge**
 - **Rate [number of publications]**
 - Direction

- Diffusion of knowledge

Immigrants are more productive researchers, even controlling for their occupational choice

$$Y_{ij} = \beta_0 + \beta_1 \text{Immigrant}_i + \beta_2 \text{Industry} + \delta_{\text{graduationyear}} + \gamma_{\text{institution}} + \sigma_{\text{department}} + \theta_{\text{year}} \quad (2)$$

	Annual Number of Publications	Number of Publications in Peer-Reviewed Journals	SNIP Weighted Number of Publications	Number of Patents
	(1)	(2)	(3)	(4)
Immigrant	0.0920*** (0.032)	0.0672*** (0.018)	0.0070** (0.003)	0.0351** (0.017)
Industry Job	-0.3939*** (0.032)	-0.2010*** (0.019)	0.0039 (0.005)	0.0100 (0.015)
Total Observations	6533	6533	6533	6533
Mean of Dep. Variable	1.6247	0.3819	0.0990	1.4317
Graduation Year FE	X	X	X	X
Year FE	X	X	X	X
PhD Institution FE	X	X	X	X
PhD Department FE	X	X	X	X

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

Notes: Estimates stem from OLS regression in which dependent variables are publication counts. SNIP weighted publications are inverse hyperbolic sine transformed publications weighted by the journal's source normalized impact per paper, a measure of a journal's reach. [b] Heteroskedastic robust standard errors, clustered at the level of the graduate, are given in parentheses.

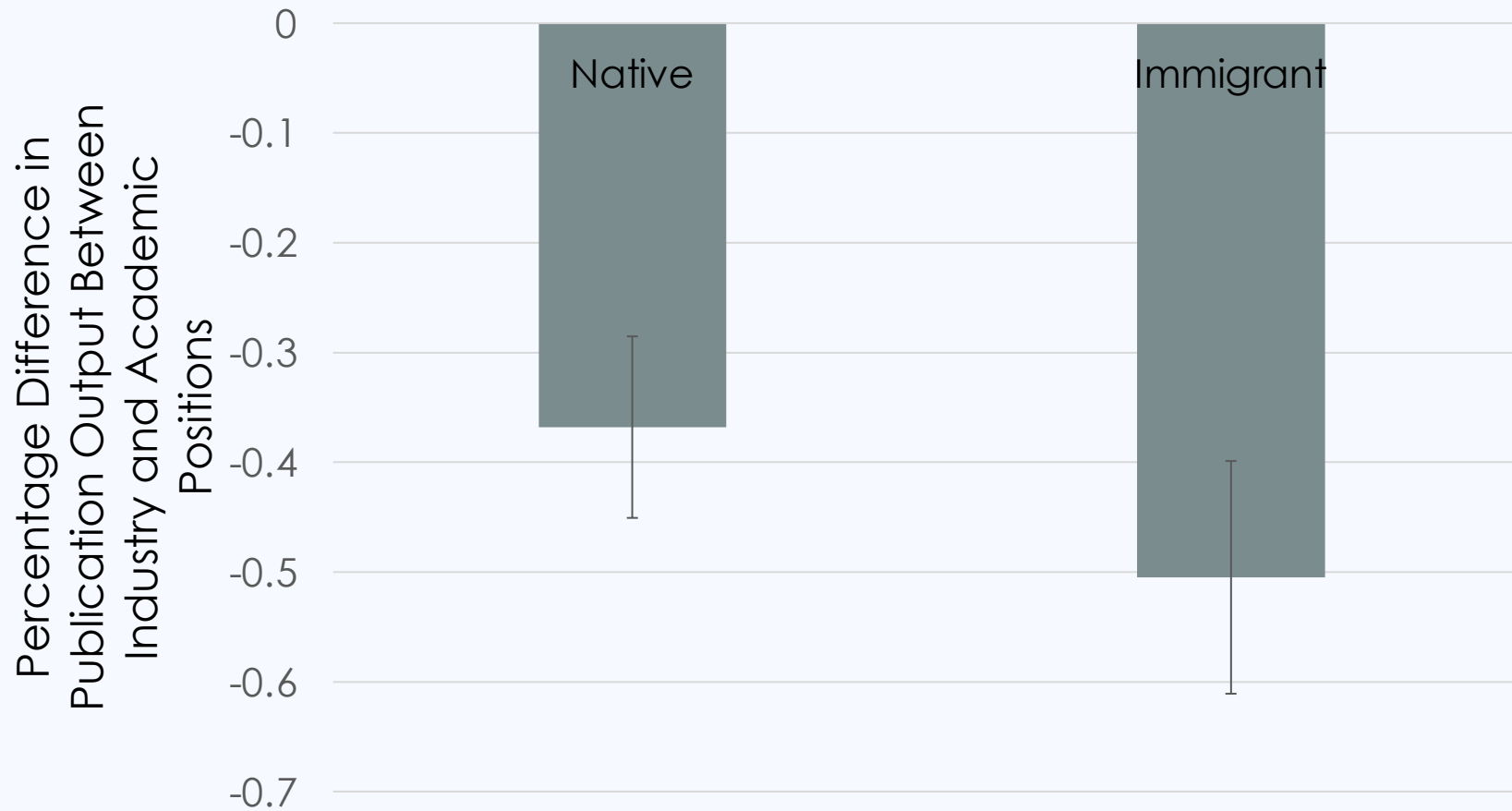
Immigrants are more productive researchers, even controlling for their occupational choice

	Annual Number of Publications	Number of Publications in Peer-Reviewed Journals	SNIP Weighted Number of Publications	Number of Patents
	(1)	(2)	(3)	(4)
Immigrant	0.0920*** (0.032)	0.0672*** (0.018)	0.0070** (0.003)	0.0351** (0.017)
Industry Job	-0.3939*** (0.032)	-0.2010*** (0.019)	0.0039 (0.005)	0.0100 (0.015)
Total Observations	6533	6533	6533	6533
Mean of Dep. Variable	1.6247	0.3819	0.0990	1.4317
Graduation Year FE	X	X	X	X
Year FE	X	X	X	X
PhD Institution FE	X	X	X	X
PhD Department FE	X	X	X	X

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

Notes: Estimates stem from OLS regression in which dependent variables are publication counts. SNIP weighted publications are inverse hyperbolic sine transformed publications weighted by the journal's source normalized impact per paper, a measure of a journal's reach. [b] Heteroskedastic robust standard errors, clustered at the level of the graduate, are given in parentheses.

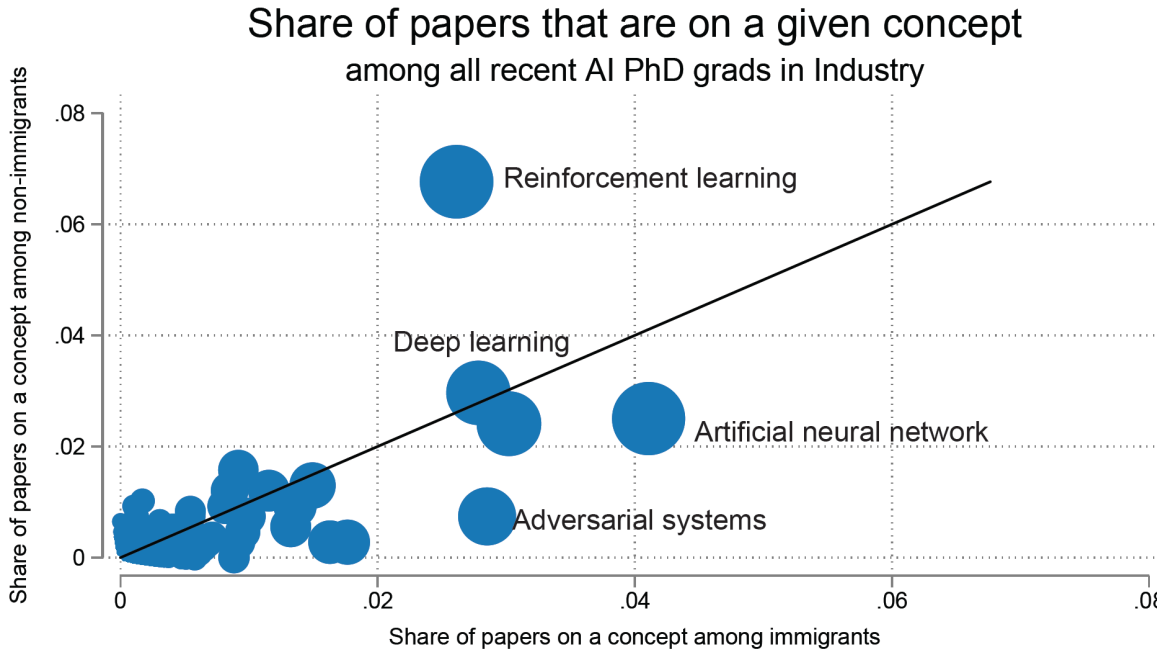
But going into industry reduces immigrant productivity in terms of publications



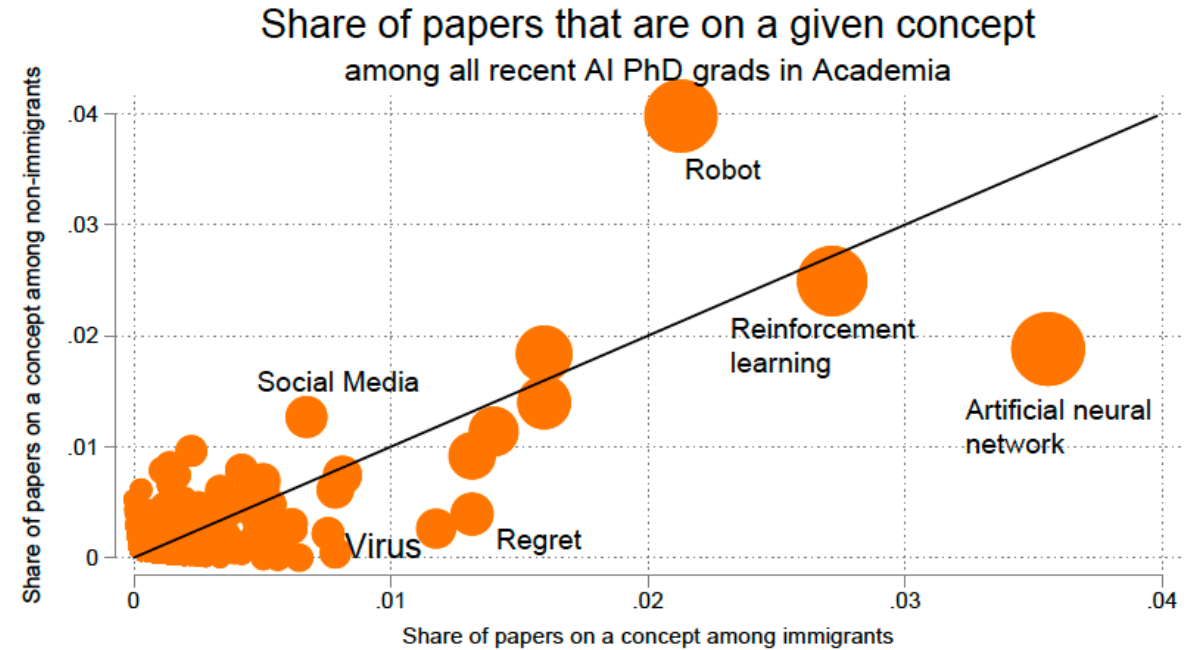
Implications of Career Choice

- **Production of knowledge**
 - Rate
 - **Direction [topics studied]**
- Diffusion of knowledge

Immigrants tend to focus on different topics than natives, regardless of their occupational choice



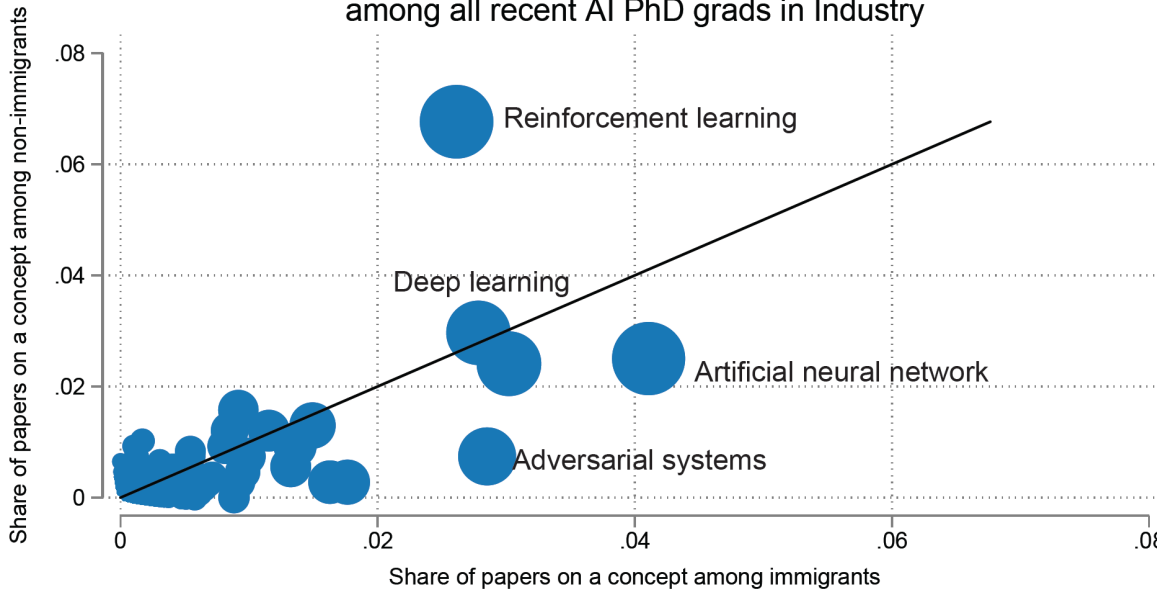
Denominator = total # of papers by all recent AI immigrant or native PhD grads in industry
 Numerator = total # of papers on a particular concept by an immigrant or native
 Most relevant concept for each paper
 Size of circle denotes the number of papers



Denominator = total # of papers by all recent AI immigrant or native PhD grads in academia
 Numerator = total # of papers on a particular concept by an immigrant or native
 Most relevant concept for each paper
 Size of circle denotes the number of papers

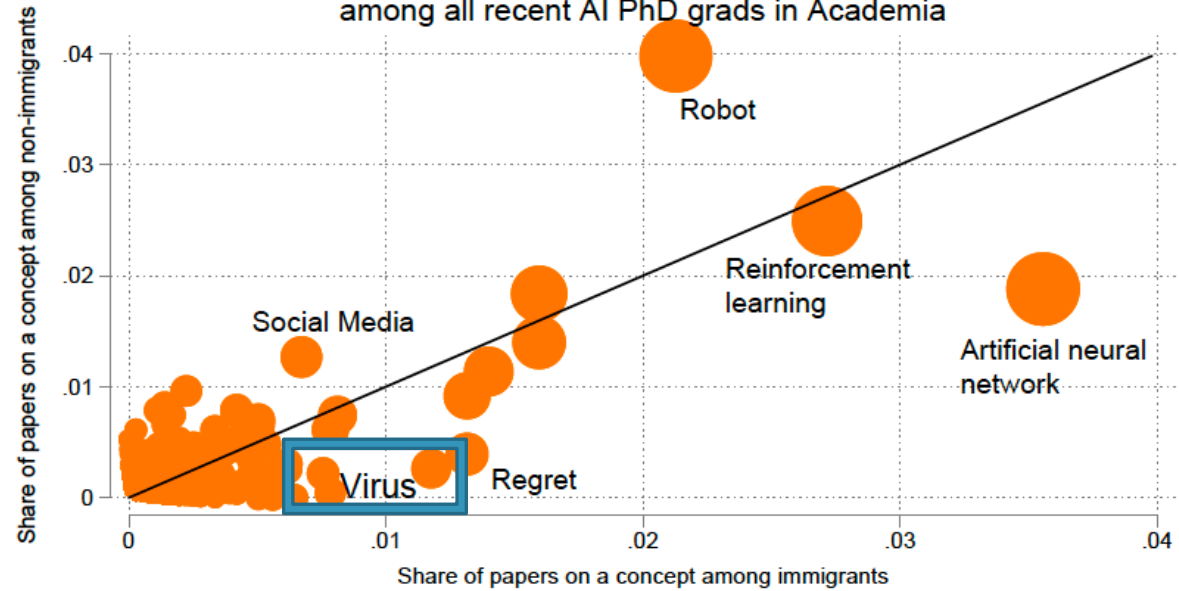
Immigrants tend to focus on different topics than natives, regardless of their occupational choice

Share of papers that are on a given concept among all recent AI PhD grads in Industry



Denominator = total # of papers by all recent AI immigrant or native PhD grads in industry
 Numerator = total # of papers on a particular concept by an immigrant or native
 Most relevant concept for each paper
 Size of circle denotes the number of papers

Share of papers that are on a given concept among all recent AI PhD grads in Academia



Denominator = total # of papers by all recent AI immigrant or native PhD grads in academia
 Numerator = total # of papers on a particular concept by an immigrant or native
 Most relevant concept for each paper
 Size of circle denotes the number of papers

PhD graduates—both immigrant and native—who go into industry have a narrower topic focus

DV = Number of concepts on publications post-graduation

	All (1)	Native (2)	Immigrant (3)
Industry	-0.0732*** (0.00877)	-0.0948*** (0.0191)	-0.0713*** (0.0106)
Immigrant	0.124*** (0.00723)		
Total Observations	9863	3350	6505
# of Pubs Control	X	X	X
PhD Department FE	X	X	X
Graduation Year	X	X	X
PhD Institution FE	X	X	X

Implications of Career Choice

- Production of knowledge
 - Rate
 - Direction [topics studied]
- **Diffusion of knowledge**

Regardless of career choice, immigrants are responsible for more (global) knowledge diffusion

$$Y_{ij} = \beta_0 + \beta_1 \text{Immigrant}_i + \beta_2 \text{Industry} + \delta_{\text{graduationyear}} + \gamma_{\text{institution}} + \sigma_{\text{department}} + \theta_{\text{year}} \quad (2)$$

	Annual Citation Count	Annual Citation Count From U.S.-Based Authors	Annual Citation Count From Non-U.S.-Based Authors
	(1)	(2)	(3)
Immigrant	0.1824*** (0.065)	0.1480*** (0.057)	0.1741*** (0.059)
Industry Job	-0.6108*** (0.064)	-0.5362*** (0.057)	-0.5072*** (0.059)
Total Observations	6533	6533	6533
Mean of Dep. Variable	1.4317	1.4317	1.4317
Graduation Year FE	X	X	X
Year FE	X	X	X
PhD Institution FE	X	X	X
PhD Department FE	X	X	X

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

Notes: [a] Estimates stem from OLS regression in which dependent variables are citation counts. [b] Heteroskedastic robust standard errors, clustered at the level of the graduate, are given in parentheses.

Regardless of career choice, immigrants are responsible for more (global) knowledge diffusion

$$Y_{ij} = \beta_0 + \beta_1 \text{Immigrant}_i + \beta_2 \text{Industry} + \delta_{\text{graduationyear}} + \gamma_{\text{institution}} + \sigma_{\text{department}} + \theta_{\text{year}} \quad (2)$$

	Annual Citation Count	Annual Citation Count From U.S.-Based Authors	Annual Citation Count From Non-U.S.-Based Authors
	(1)	(2)	(3)
Immigrant	0.1824*** (0.065)	0.1480*** (0.057)	0.1741*** (0.059)
Industry Job	-0.6108*** (0.064)	-0.5362*** (0.057)	-0.5072*** (0.059)
Total Observations	6533	6533	6533
Mean of Dep. Variable	1.4317	1.4317	1.4317
Graduation Year FE	X	X	X
Year FE	X	X	X
PhD Institution FE	X	X	X
PhD Department FE	X	X	X

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

Notes: [a] Estimates stem from OLS regression in which dependent variables are citation counts. [b] Heteroskedastic robust standard errors, clustered at the level of the graduate, are given in parentheses.

But career choice shapes the geographical diffusion of knowledge. Per publication, immigrants in academia are more likely to be cited globally (and in industry less likely to be cited locally)

Industry

	Pub Citations From U.S. Based Authors (1)	Pub Citations From Non-U.S. Based Authors (2)
Immigrant	-0.2905** (0.114)	-0.0652 (0.120)
Total Observations	4298	4298
Mean of Dep. Variable	10.6764	17.3941
Year FE	X	X
PhD Department FE	X	X

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

Notes: [a] Estimates stem from OLS fixed effects regression in which dependent variables are inverse hyperbolic sine counts of citations by different types of authors. [b] Heteroskedastic robust standard errors, clustered at the level of the graduate, are given in parentheses.

Academia

	Pub Citations From U.S. Based Authors (1)	Pub Citations From Non-U.S. Based Authors (2)
Immigrant	0.1037 (0.109)	0.2202** (0.105)
Total Observations	5881	5881
Mean of Dep. Variable	9.1248	13.6242
Year FE	X	X
PhD Department FE	X	X

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

Notes: [a] Estimates stem from OLS fixed effects regression in which dependent variables are inverse hyperbolic sine counts of citations by different types of authors. [b] Heteroskedastic robust standard errors, clustered at the level of the graduate, are given in parentheses.

Conclusion & Implications

Immigrant graduates of AI PhD programs are much more likely than non-immigrant graduates to go into industry

- Suggestive evidence that this is a **choice**
 - Persists when controlling for observable measures of ability and selection
 - The opposite result from what we'd expect if visa constraints dominated

Conclusion & Implications

Immigrant graduates of AI PhD programs are much more likely than non-immigrant graduates to go into industry

- Suggestive evidence that this is a **choice**
 - Persists when controlling for observable measures of ability and selection
 - The opposite result from what we'd expect if visa constraints dominated

This choice has significant implications for the shape of science

- **Rate:** Fewer publications by some of the top graduates
 - Either fewer advances in science, or less public dissemination of scientific advances
- **Direction:** Different topic focus in industry vs academia + Different topic focus for immigrants vs natives
 - “missing knowledge”?
- **Diffusion:** A decline in citations outside the US
 - less overall diffusion of knowledge globally