The Role of Immigrants and Foreign Students
in Science, Innovation, and Entrepreneurship
Cambridge MA, January 26, 2018

Returnee migrants' self-selection: Evidence for Indian inventors

Stefano Breschi Francesco Lissoni Ernest Miguelez

Work in Progress

Objective

- Highly skilled return migration: under-studied area → data shortage
- Explore potentialities and limitations of professional social network data (LinkedIn)
- Focus on Indian migrant inventors in ICT

Database: Indian Inventors in ICT

1. Inventors from USPTO patents assigned to 178 largest(*) US public firms in ICT, 1975-2016
   - *Sources*: Patentsview / Compustat ; *Obs:* 262,849

   (*) >250 granted patents 1975-2016



2. Identification of Indian(-origin) inventors, by name & surname
   - *Sources*: IBM-GNR / Breschi et al. (2017)



3. Match to LinkedIn profiles (*by* name & surname of inventors/LinkedIn members *plus* name of assignees/employers as per LinkedIn profiles)
   - *Source*: LinkedIn (June 2016)

Q: Why LinkedIn? → A: Cross-border information on education and employment

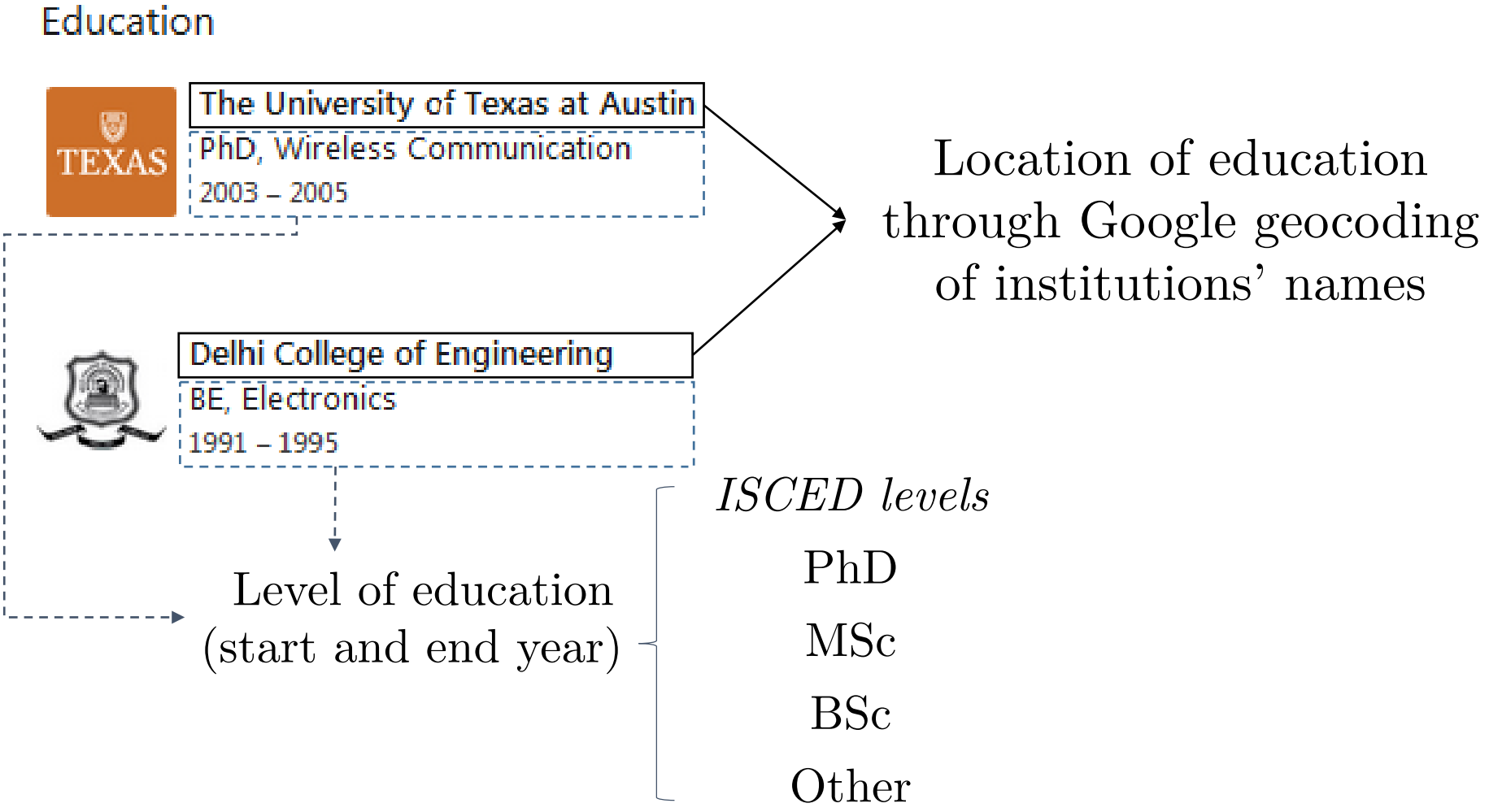
- Less complete and more prone to error when compared employer-employee data (as e.g. in Rooth-Sarela, 2014), but spanning across countries
- Information on education not just at destination (e.g. NSCG), but also at origin
- Info on undergraduate education may replace/improve upon country of birth when it comes to define migration

Indian Inventors in ICT: LinkedIn matching outcome

	Incomplete profile	Complete profile*	No match	Total
All inventors	1789	9050	13178	24017
%	*7.4*	*37.7*	*54.9*	*100.0*

* Information on both employment and education

Coding *education* information



Coding *employment* information



$1\{0\}$: if inventor *self-reports* **location** of **employment** in India

$1\{0\}$: if inventor *self-reports* **location** of **employment** *abroad* (i.e. any country but India)

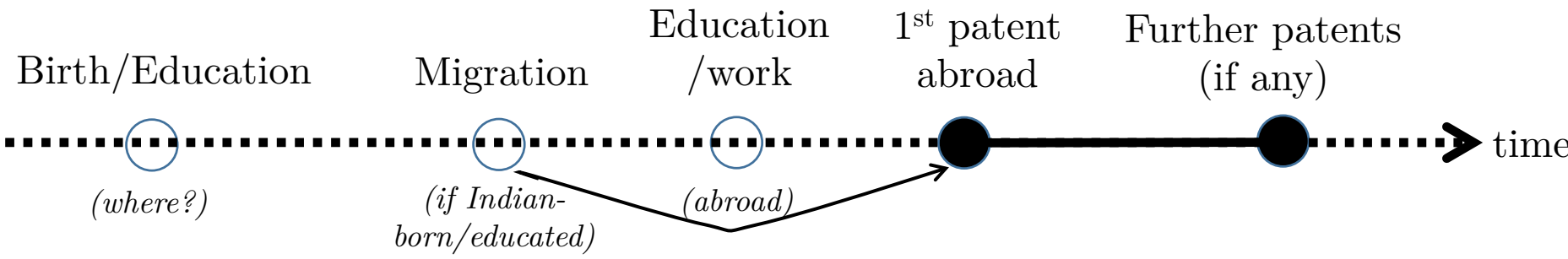
Start / end year of employment spell

Patent- **vs** *education/employment-***based definition of migrant inventor**

Patent-based	Indian-named (or Indian national) inventor with at least one patent with non-Indian address
Education/employment-based	Indian-named (or Indian national) inventor with at least some education or work experience in India before: - patenting in the US (abroad) *or* - being employed in the US (abroad)

Patent-based definition of migrant inventor

- Widespread use: Kerr (2008), Agrawal et al. (2011), Breschi et al. (2017)
- Adopted due to data constraints
- Left truncation:
  - **Type I error** (false positives: Indian diaspora; non-Indian migrants) → it turns out to be possibly large: must be corrected (and we can)
  - **Type II errors** (false negatives: migrants non-patenting abroad) → not so large (we can correct anyway)
  - **Heterogeneity of migration channels/motives**: work (e.g. H-1B visa) vs education (graduate) → pretty large (we can measure it)



Patent-based migrants vs non-migrants

		Nr	%
At least one patent abroad?	Yes	7035	*77.7*
	No	2015	*22.3*
Total		9050	*100*

Type I error

Any education/employment*/patent in India?

No		Yes		Total
Nr	%	Nr	%	Nr
1,473	*20.9*	5,562	*79.1*	7,035
FALSE POSITIVES (diaspora / 2nd generation)		**TRUE POSITIVES**		

* As self-reported in resume

Patent-based migrants vs non-migrants

		Nr	%
At least one patent abroad?	Yes	7035	*77.7*
	No	2015	*22.3*
Total		9050	*100*

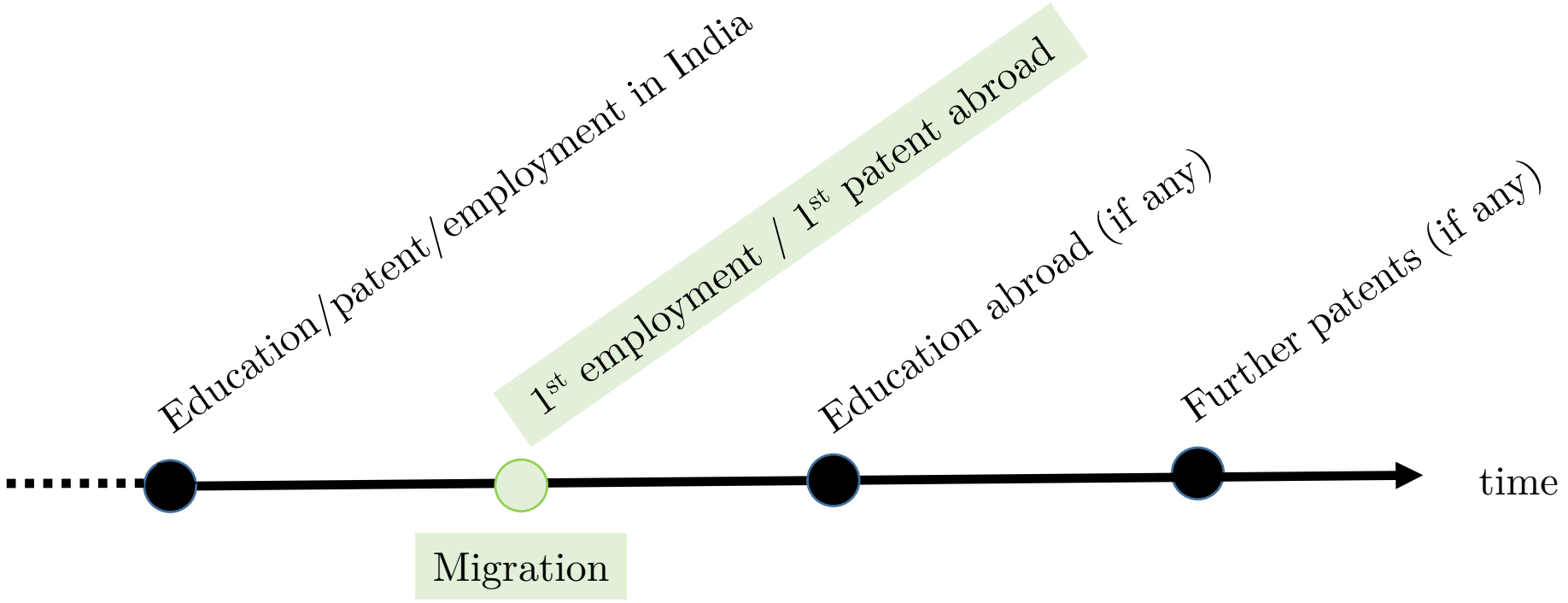
Type II error

Any education/employment* abroad?

No		Yes		Total
Nr	%	Nr	%	Nr
1,618	*80.3*	397	*19.7*	2,015
NON MIGRANTS		**FALSE NEGATIVES**		

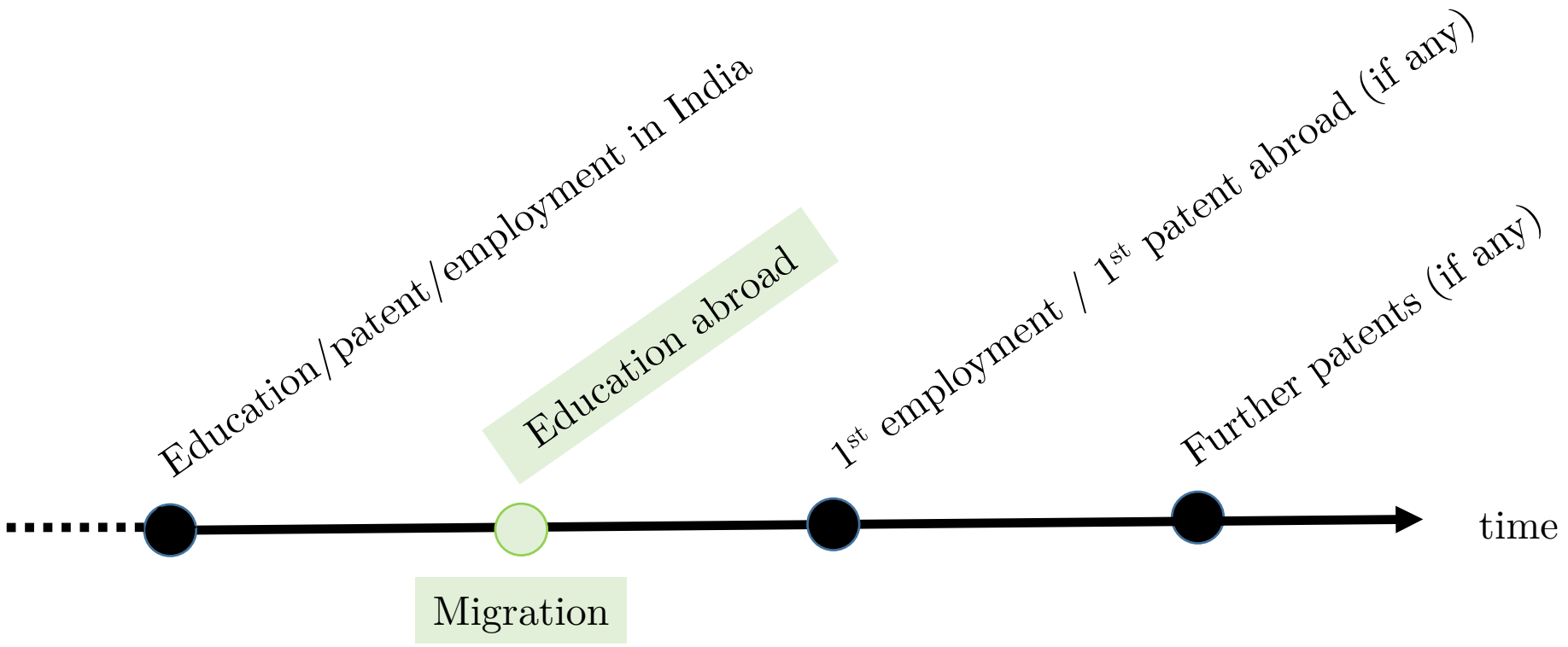
* As self-reported in resume

From *patent-based* to *education/employment*-based definition of migrant inventor



Work migration channel

From *patent-based* to *education/employment*-based definition of migrant inventor



Education migration channel

Final sample

	Migration channel		
	Education	**Work**	**Total**
True positives	3,826	1,736	5,562
%	*64.2*	*29.1*	*93.3*
False negatives	382	15	397
%	*6.4*	*0.3*	*6.7*
Total	4,208	1,751	5,959
%	*70.6*	*29.4*	*100.0*

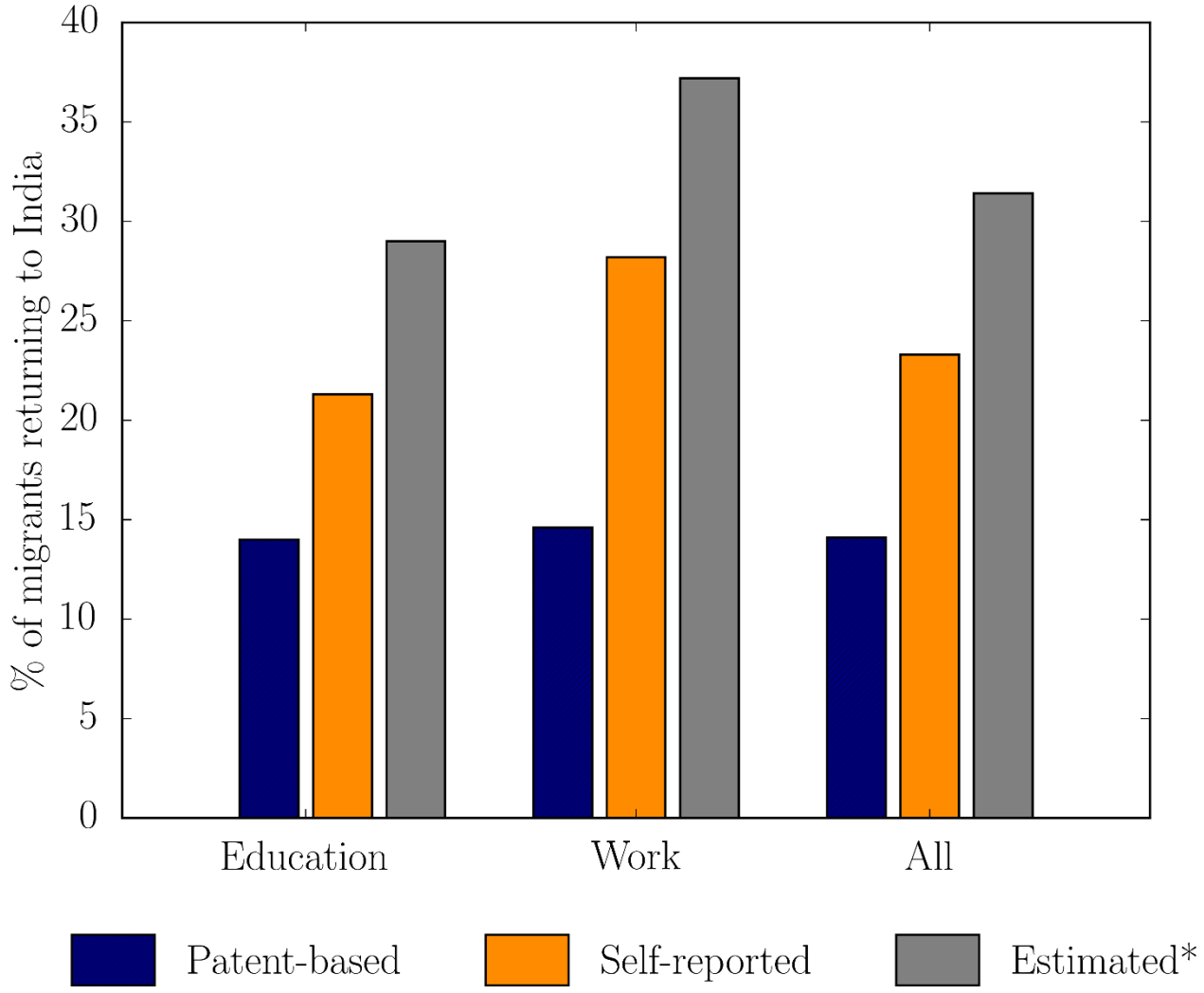
Cohorts of migrants (migration year) by migration channel

	1960s	**1970s**	**1980s**	**1990s**	**2000s**	**2010s**	**Total**
Education	28	152	817	1,822	1,314	75	4,208
	100%	*97%*	*95%*	*84%*	*54%*	*23%*	*71%*
Work	0	4	40	351	1,101	255	1,751
	0%	*3%*	*5%*	*16%*	*46%*	*77%*	*29%*
Total	28	156	857	2,173	2,415	330	5,959

Patent- **vs** *education/employment-*based definition of **return** migrant inventor

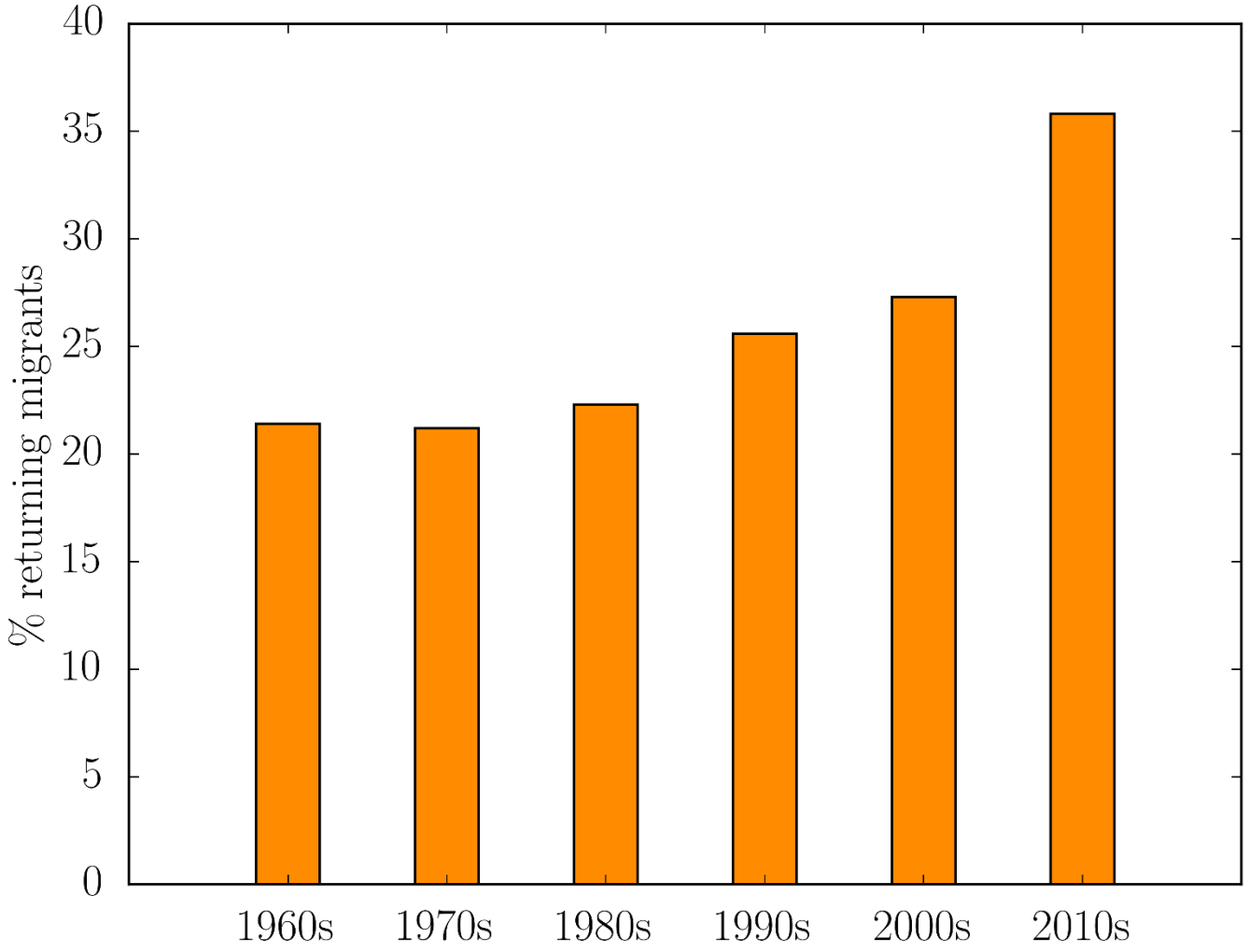
Patent-based	The migrant inventor patents in India after patenting in the US
Self-reported	The migrant inventor patent *and/or declares to be* employed in India after patenting in the US
Estimated	The migrant inventor patent *and/or* is employed in a *company with many Indian employees*, after patenting in the US

Return migration rates (% of migrants), by migration channel



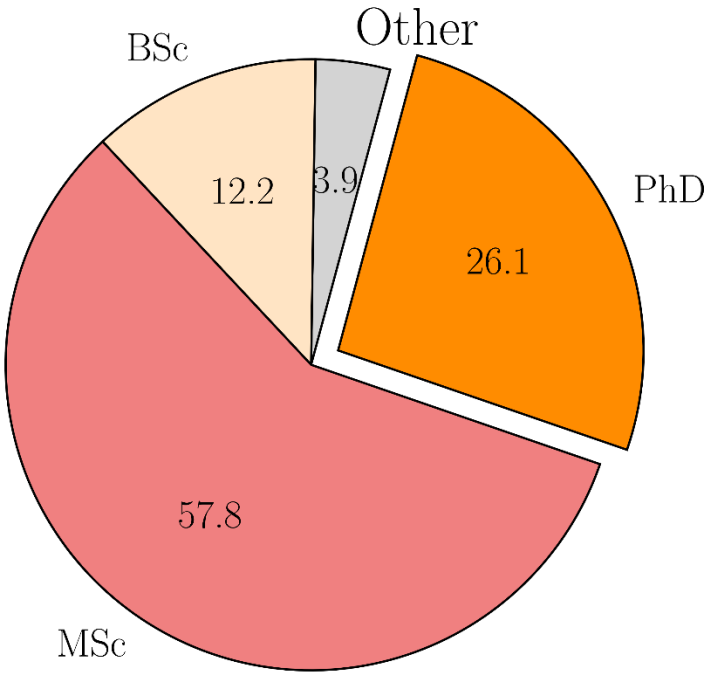
* Estimated: post-migration employer has no less than 50% employees in India

% return rate (job in India*) by migration cohort (all channels)

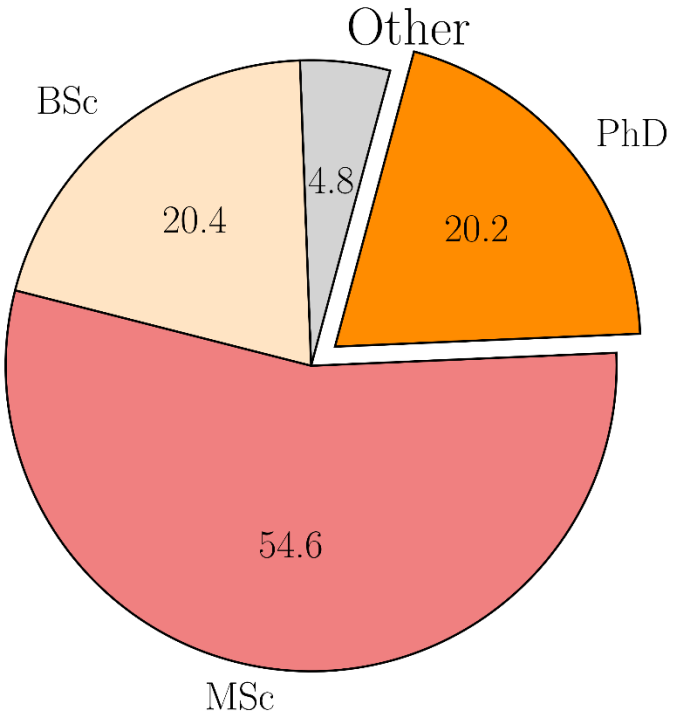


*As self- reported in the resume

Returning* vs Permanent migrants
% distribution by highest degree



Permanent migrants

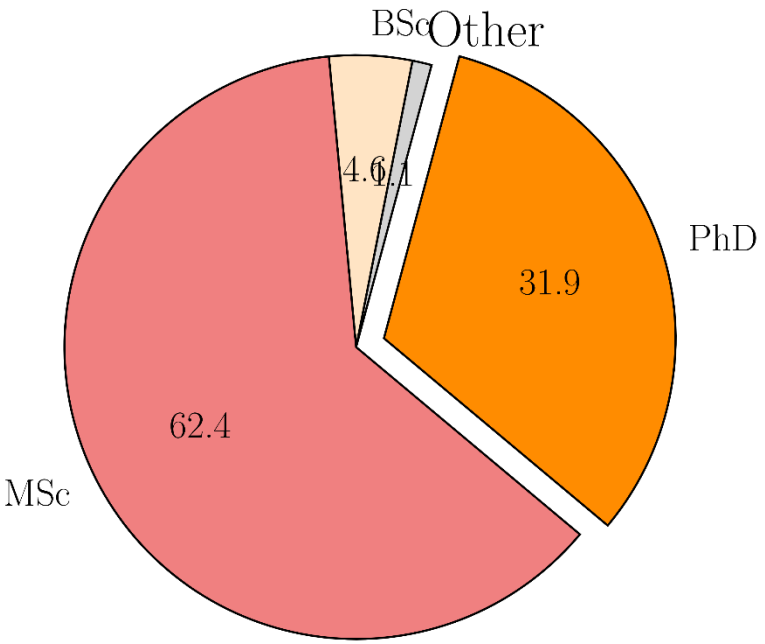


Returning migrants

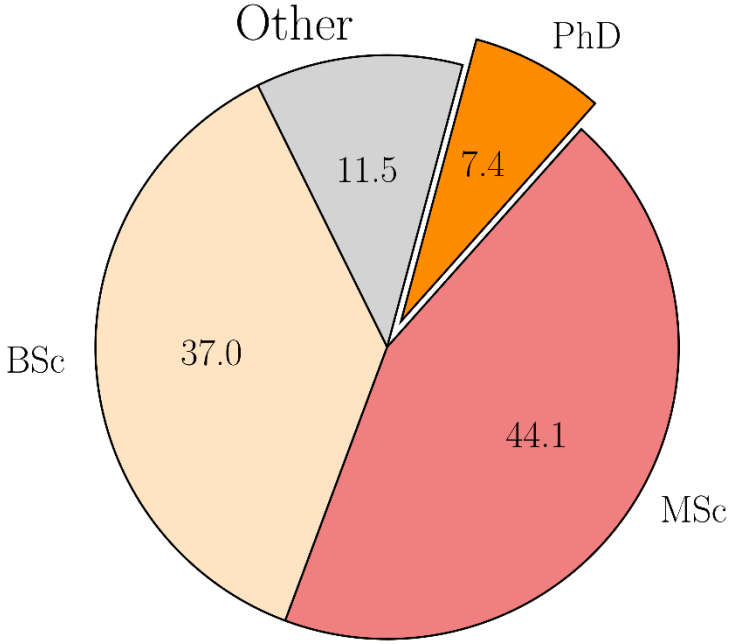
* Returnees based on self-reported definition

Migrants education by migration channel

% distribution by highest degree

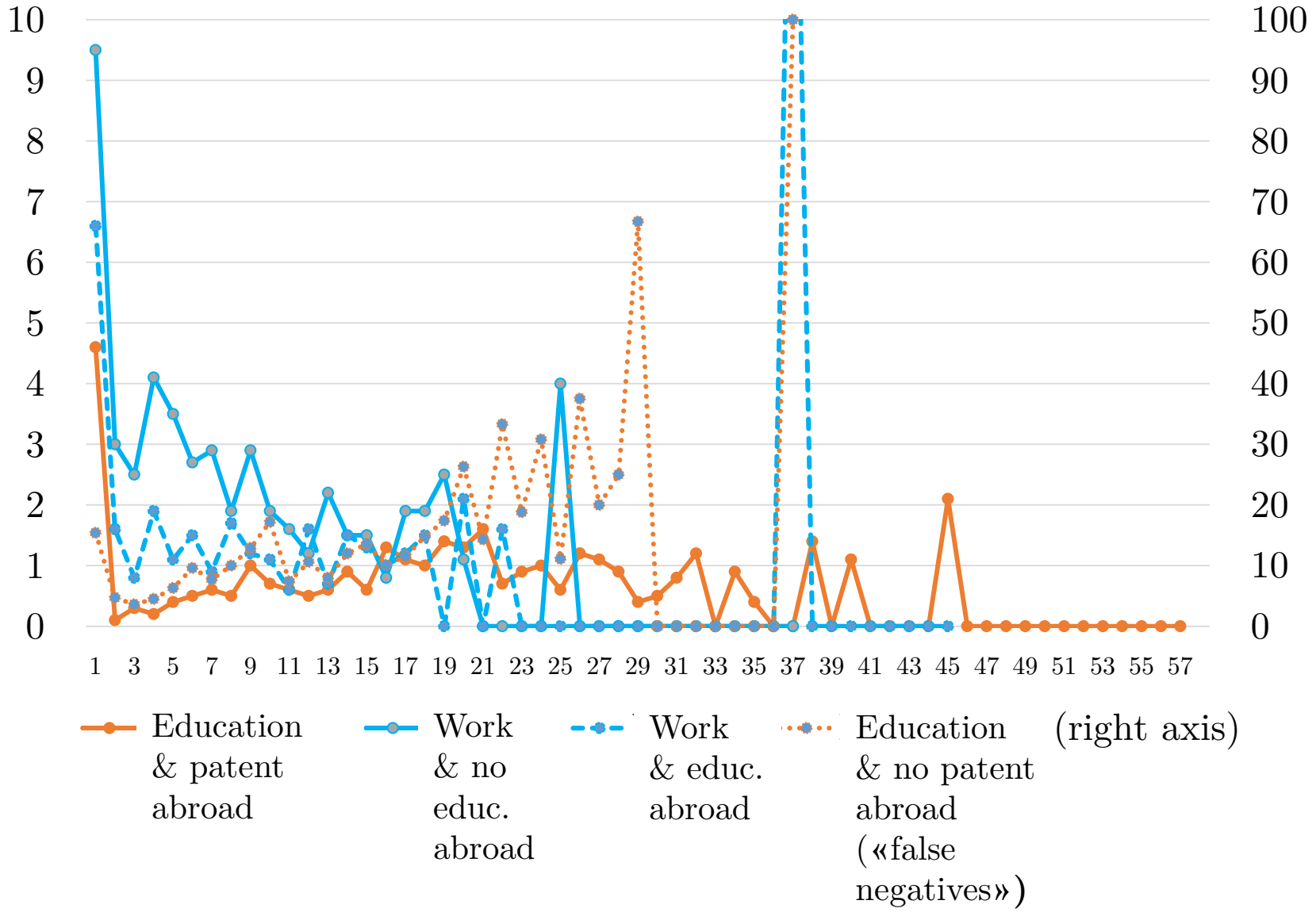


Education channel



Work channel

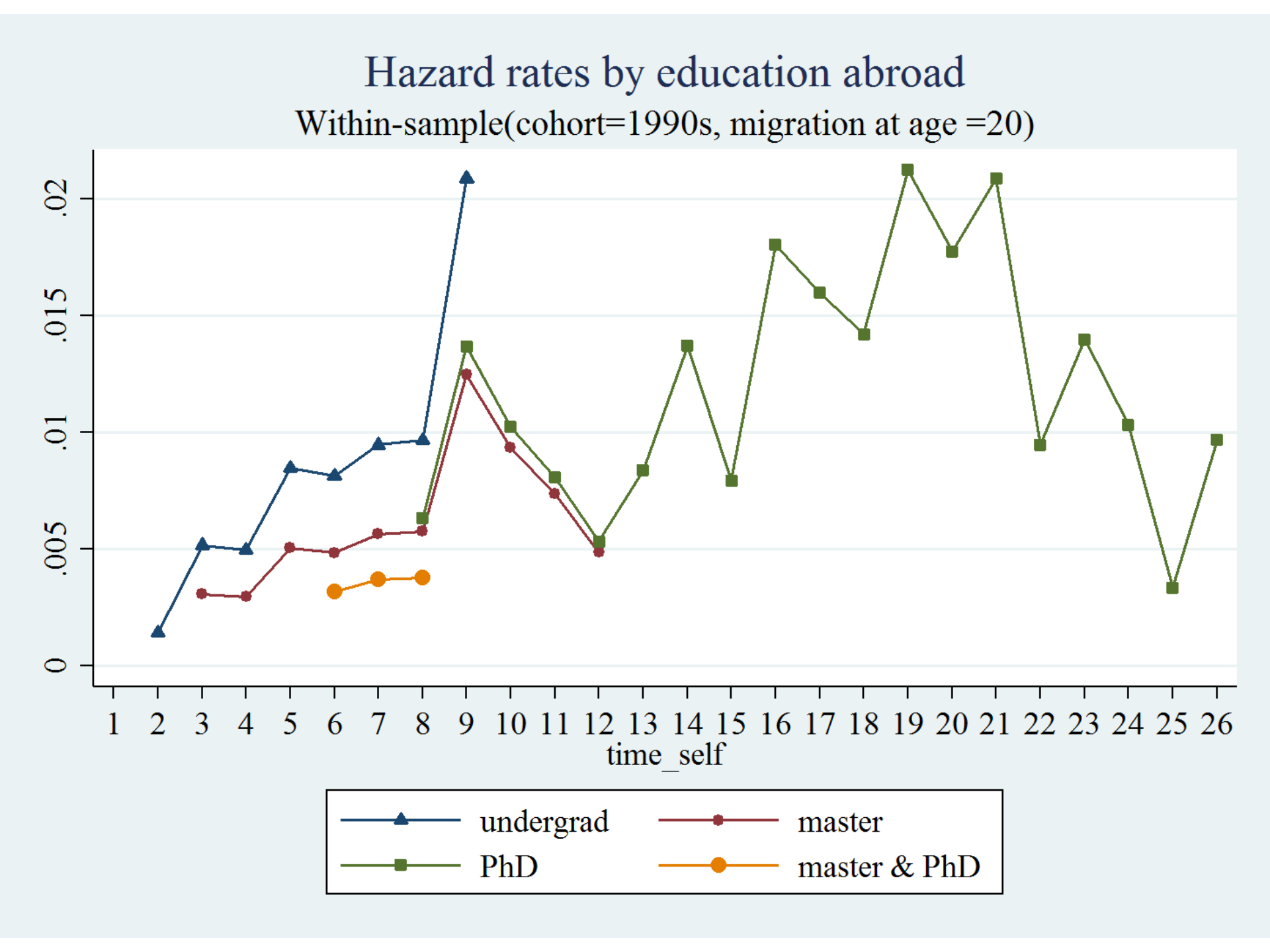
Per-year % return rate (job in India): By Migration Channel and Subgroup



Event history analysis of return risk (discrete time, semi-parametric proportional hazard) – « False negatives » excluded

	All cohorts			Recent cohorts (since 1990)		
	All migrants	***All migrants***	***Only Educat.***	***All migrants***	***Only Educat.***	***Only Work***
Age at migration	-0.03***	-0.06***	-0.03***	-0.09***	-0.07***	-0.09***
Master at migration	0.25**	0.39***	0.37**	0.61***	0.72***	0.04
PhD at migration	-0.08	0.44*		0.42		1.13**
Migr. channel: Education	-0.89***	-0.67***		-0.75***		
Student abroad	-0.41***	-0.35***	-0.44***	-0.34***	-0.27	-1.17***
Master (while abroad)	-0.28***	-0.37***	-0.48***	-0.42***	-0.51***	0.10
PhD (while abroad)	-0.28***	-0.28**	-0.42***	-0.31**	-0.43***	-0.48
MBA (while abroad)	0.12	0.09	-0.13	0.28**	0.12	0.38**
Patent stock at migration	0.16***	0.15***	0.84***	0.17***	0.81***	0.17***
Cum. patent stock abroad	0.006***	0.005**	0.003	0.007***	0.005*	0.014***
Time dummies	Y	Y	Y	Y	Y	Y
Cohort dummies	N	Y	Y	Y	Y	Y
Observations	*95,170*	*95,170*	*75,364*	*67,671*	*49,069*	*18,375*
Chi2	*20685*	*19958*	*13074*	*15907*	*9439*	*5675*
LogL	*-5943*	*-5851*	*-3582*	*-4897*	*-2700*	*-2109*

Robust standard errors ; *** p<0.01, ** p<0.05, * p<0.1


Hazard rates by education abroad
Within-sample(cohort=1990s, migration at age =20)
0
.005
.01
.015
.02
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
time_self
undergrad
master
PhD
master & PhD

Preliminary results

- Some evidence of education-based negative self-selection conditional on migration channel (education, MSC and PhD)
- Positive self-selection on MBA for work channel
- Opposite evidence on patent-based selection

Further steps

- US vs other countries of destination
- Revise time-scale for duration analysis → grouping years
- University ranking in India and the US
  - India: IIT vs other institutions
  - US: Carnegie classification
- Weighting patents by forward citations

Indian Institute of Technology
% of inventors that attained degree from IIT

	PhD	%	MSc	%
IIT*	123	80.4	932	46.8
Other	30	19.6	1,059	53.2
Total	153	100.0	1991	100.0

*Include Indian Institute of Science