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

Return migrants' self-selection: Evidence for Indian inventors

Cambridge MA, April 27, 2018

Stefano Breschi   **Francesco Lissoni**   **Ernest Miguelez**

Bocconi
Università Commerciale
Luigi Bocconi

université
de BORDEAUX

Motivation & Background

- Historical importance of return migration → 20-50% adults return after 5 years, but large heterogeneity across countries (OECD, 2008)

- Classic topic in migration economics: opposite self-selection upon arrival and return (Borjas and Bratsberg, 1996; also Dustman and Görlach, 2016)

- BUT limited evidence in general and especially on highly skilled (see Kerr, 2017; few exceptions: Gaulé, 2014; Kahn and MacGarvie, 2012)

- Special interest for STEM returnees → key role in knowledge diffusion and entrepreneurship (Kahn and MacGarvie, 2012; Jonkers and Cruz-Castro, 2013; Gibson and McKenzie, 2014; Nanda and Khanna, 2010; Choudhury, 2016)

Objective

1) Contribute to fill the data gap for a specific category of STEM migrants, highly exposed to return risk: Indian inventors in the US of ICT companies
   ➢ Heavy users of H1B visas
   ➢ Student visas to enter US universities

2) Positive/negative self-selection
   a. With respect to education and (unobservable) skills
   b. Different migration channels (diff. motivations to migrate) → different results?

3) Explore potentialities and limitations of professional social network data (LinkedIn) + link with patent data

Data: Indian Inventors in ICT

1. Inventors from USPTO patents assigned to 179 largest[*] US public firms in ICT, 1975-2016
   - *Sources*: Patentsview / Compustat ; *Obs:* 262,847

   [*] *>250 granted patents 1975-2016*

2. Identification of Indian(-origin) inventors, by name & surname → **24,017 individuals**
   - *Sources*: IBM-GNR / Breschi et al. (2017)

3. Match to LinkedIn profiles (name of inventors + name of assignees/employers as per LinkedIn profiles) → **10,839** individuals (**8,982** with either educ. level or birth year)
   - *Source*: LinkedIn (June 2016)

Data: Indian Inventors in ICT

Distribution of LinkedIn matched and unmatched inventors by application year of the first patent at the USPTO
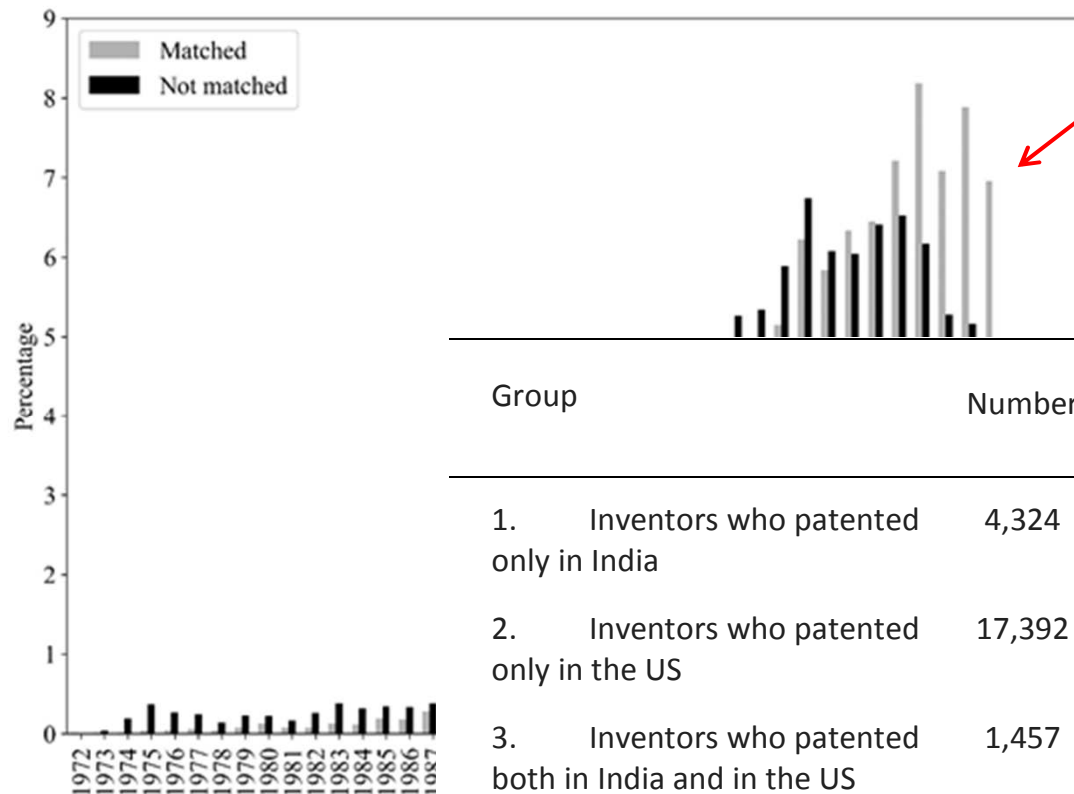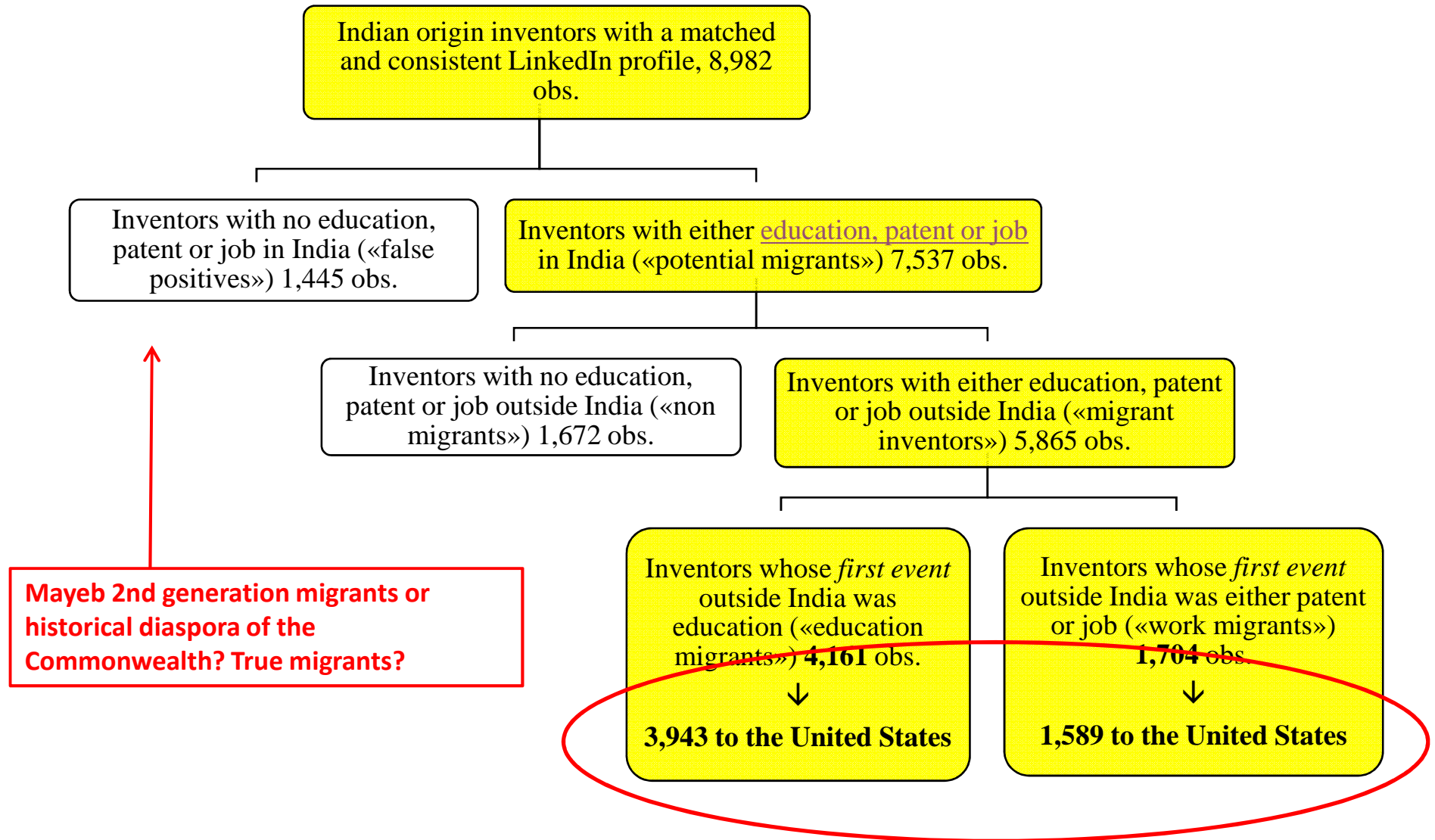

More matches for younger cohorts

Group	Number	Number with a matched LinkedIn profile	% with a matched LinkedIn profile
1. Inventors who patented only in India	4,324	2,003	46.3
2. Inventors who patented only in the US	17,392	6,088	35.0
3. Inventors who patented both in India and in the US	1,457	593	40.7
4. Others	844	298	35.3
All Indian inventors	24,017	8,982	37.4

Differences statistically significant
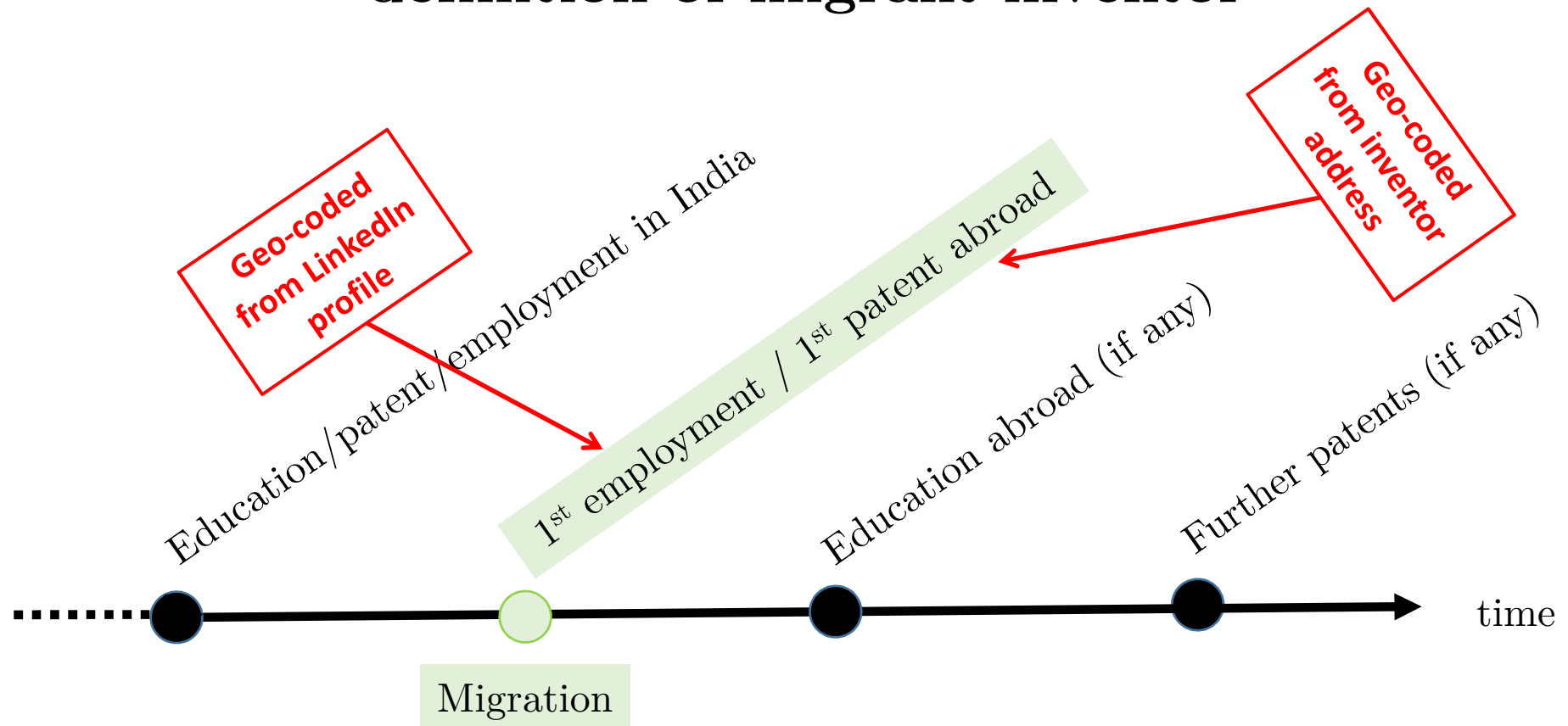
Data: Indian Inventors in ICT

Indian origin inventors with a matched and consistent LinkedIn profile, 8,982 obs.

Inventors with no education, patent or job in India («false positives») 1,445 obs.

Inventors with either education, patent or job in India («potential migrants») 7,537 obs.

Inventors with no education, patent or job outside India («non migrants») 1,672 obs.

Inventors with either education, patent or job outside India («migrant inventors») 5,865 obs.

Mayeb 2nd generation migrants or historical diaspora of the Commonwealth? True migrants?

Inventors whose *first event* outside India was education («education migrants») 4,161 obs.

↓

3,943 to the United States

Inventors whose *first event* outside India was either patent or job («work migrants») 1,704 obs.

↓

1,589 to the United States

Data: Indian Inventors in ICT

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

Geo-coded from LinkedIn profile

Education/patent/employment in India

Education abroad

1ˢᵗ employment / 1ˢᵗ patent abroad (if any)

Further patents (if any)

Migration

time
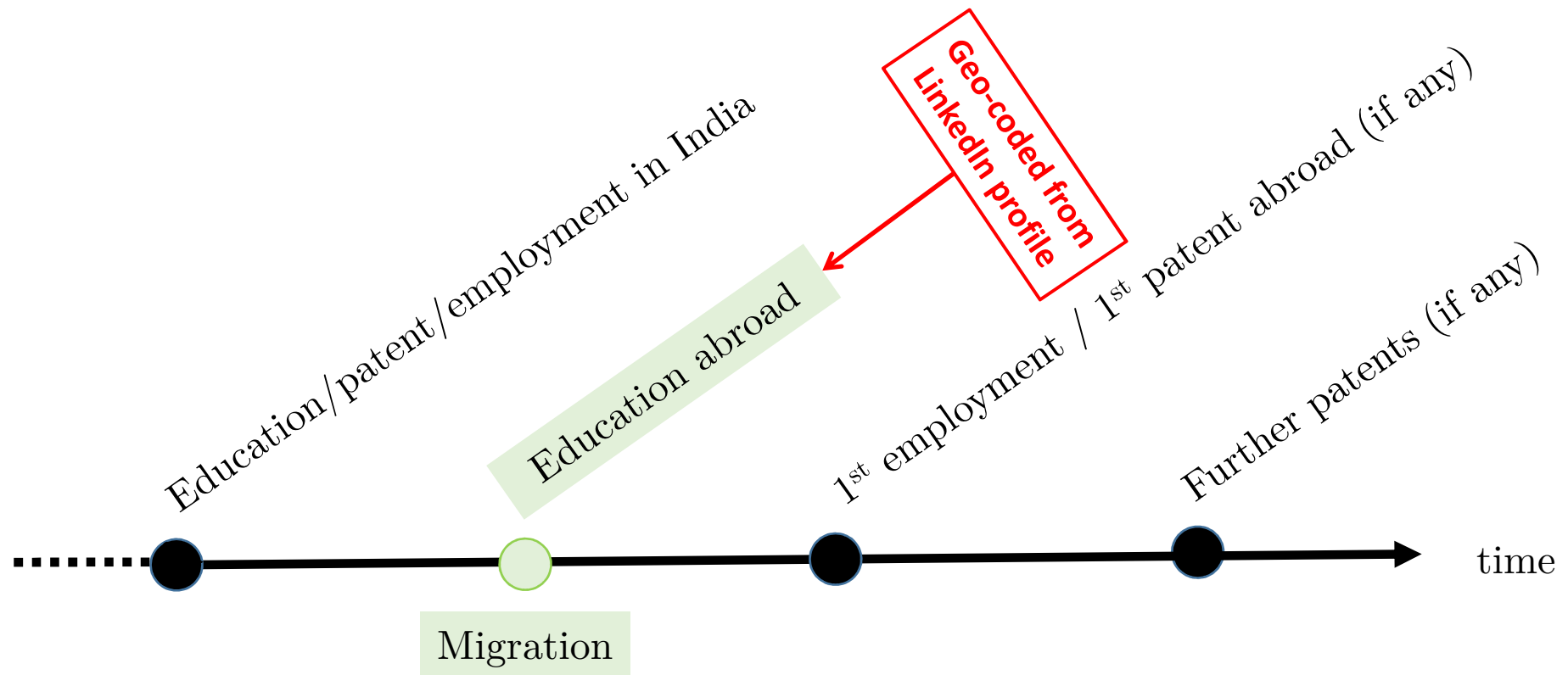
Education migration channel

Data: Indian Inventors in ICT

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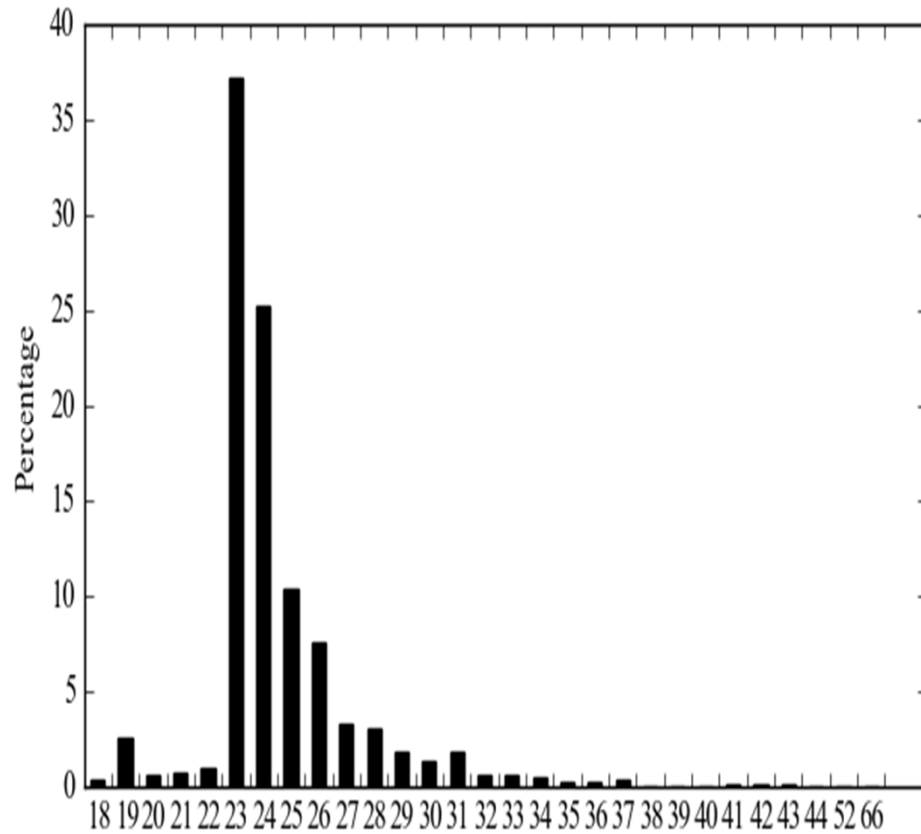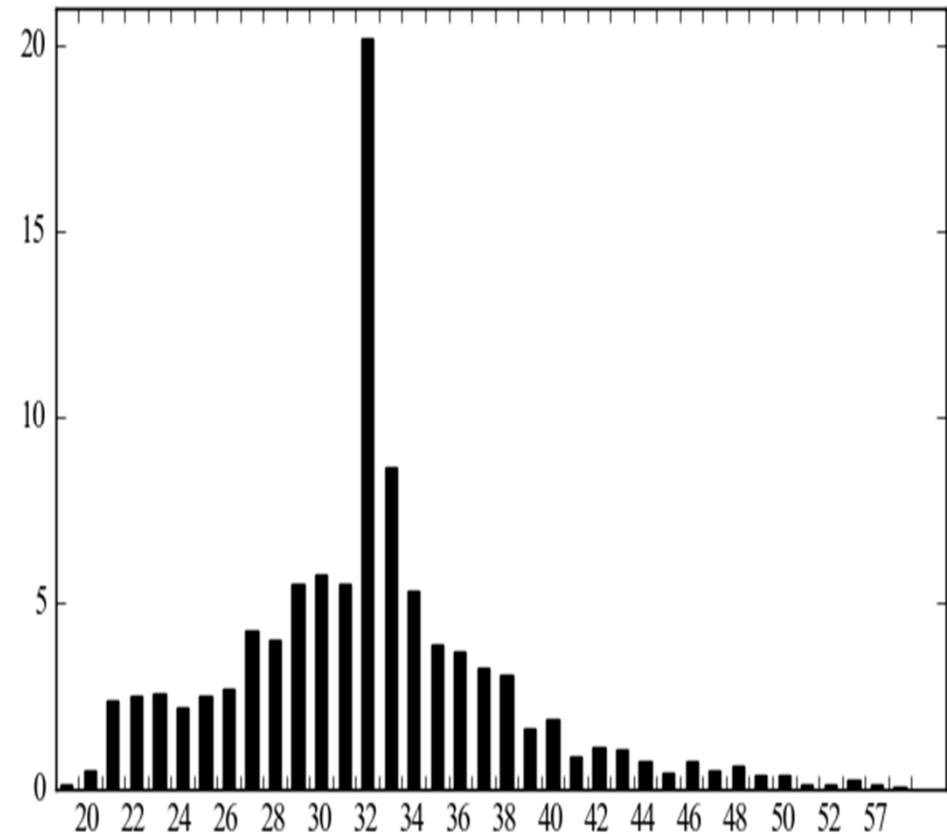

Work migration channel

Estimated age at migration (percentage distribution of all education migrants to the US)
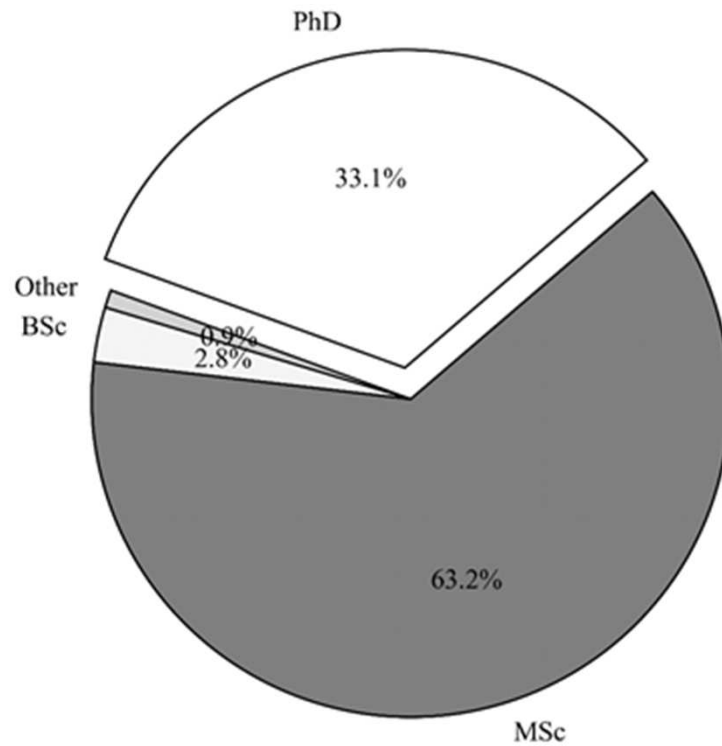
Education channel

Work migrants

Note: Age estimated from education info from LinkedIn (or average age of same cohort inventors if former was missing)
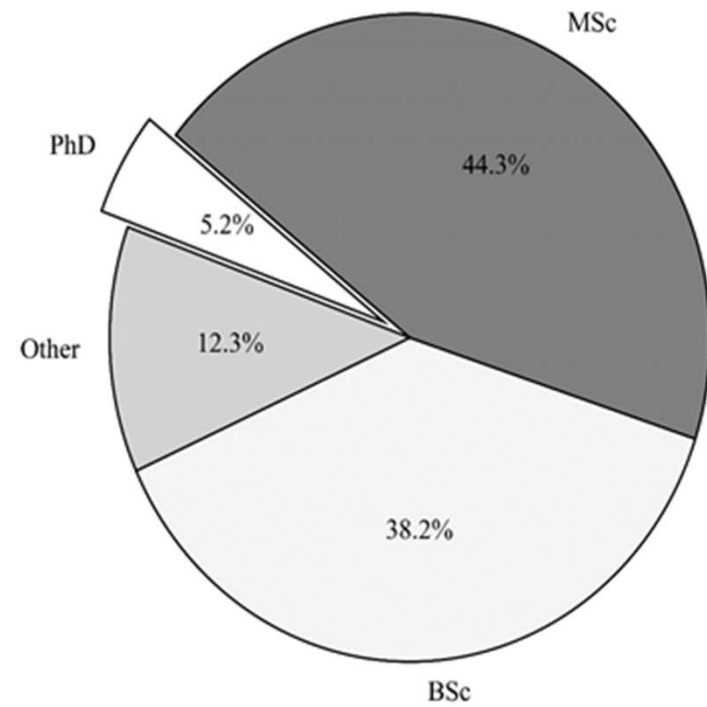
Highest educational attainment, percentage distribution

Education migrants

Work migrants

Migrants to the US by cohort and channel

4362 individuals

Channel	1960	1970	1980	1990	2000	2010	Total
Education	19	102	697	1739	1315	71	3943
%^{column}	100	95.3	95.2	85.9	56.3	22.8	71.3
Work	0	5	35	286	1022	241	1589
%^{column}	0.0	4.7	4.8	14.1	43.7	77.2	28.7
All channels	19	107	732	2025	2337	312	5532
%^{column}	100	100	100	100	100	100	100
%^{row}	0.3	1.9	13.2	36.6	42.2	5.6	100

How to measure return migration?

P*atent-* **vs** *education/job-***based definition**

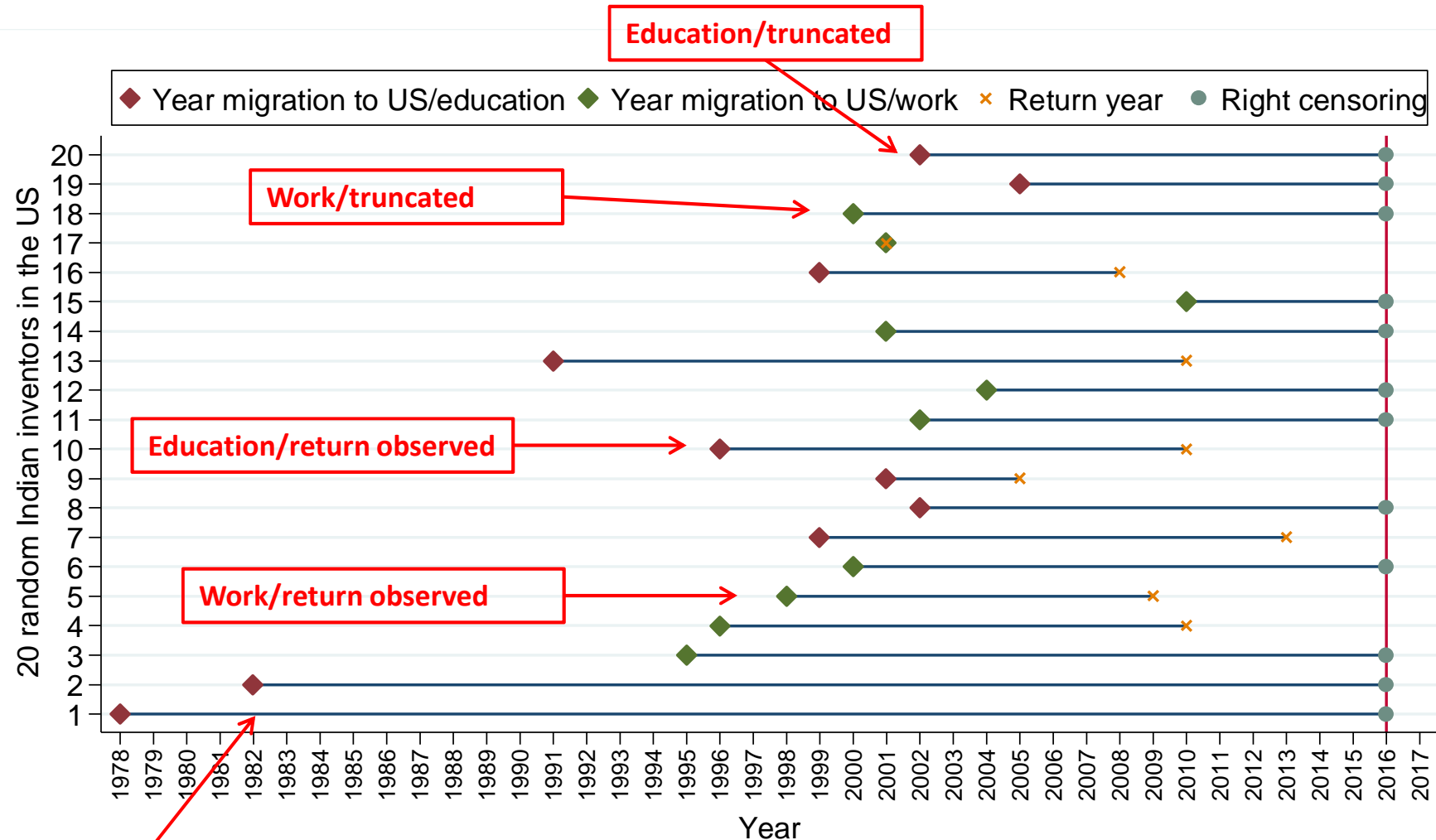|	
Patent-based	The migrant inventor patents in India after patenting in the US (e.g., Oettl and Agrawal, 2008; Agrawal et al., 2011; Breschi et al, 2017)
Job & Patent	The migrant inventor patent *and/or declares to be* employed in India after patenting in the US → higher recall

Total return rates by migration channel
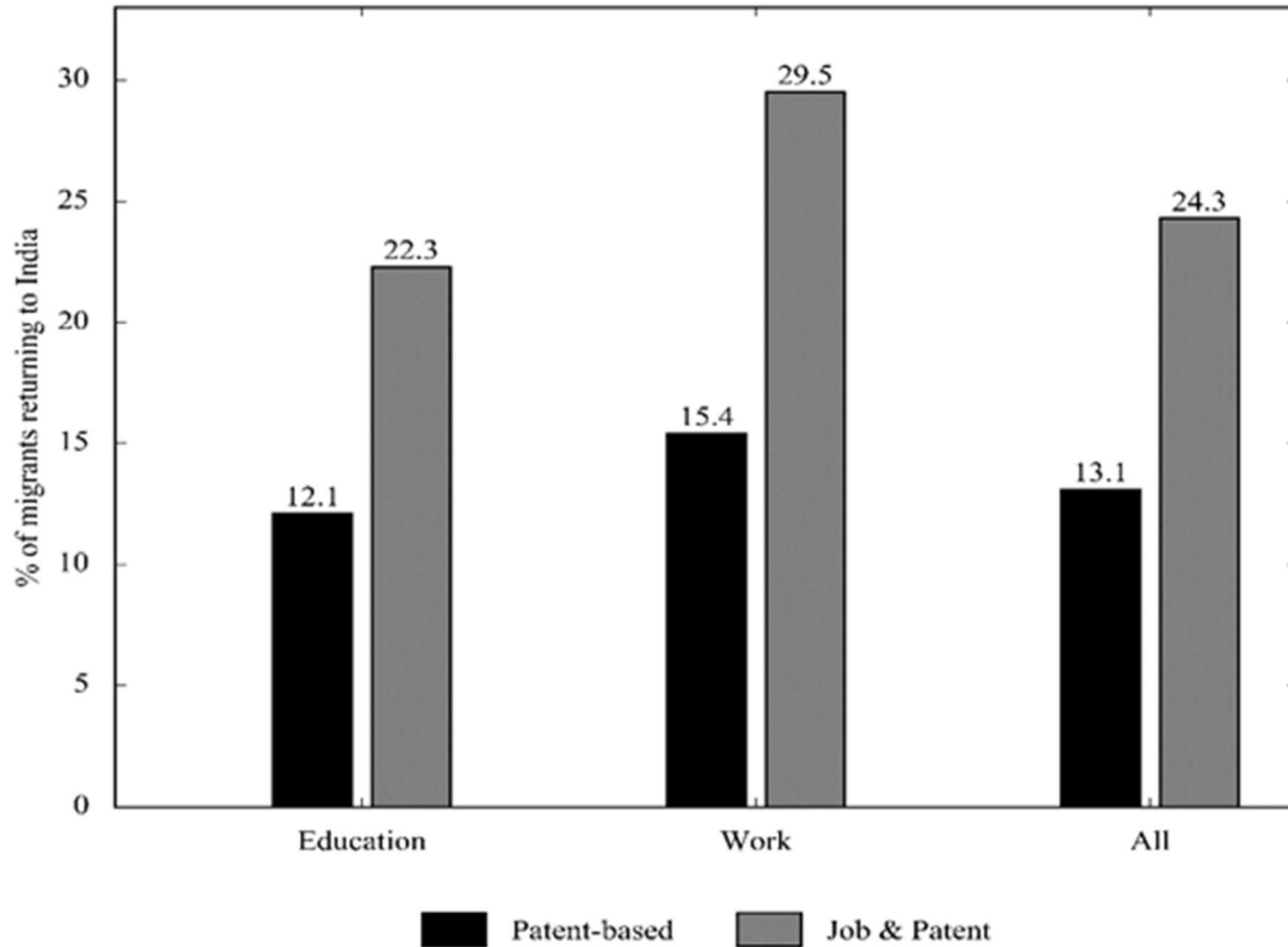
Data structure: return, right truncation, and length of stay

Survival function, by migration channel

Kaplan-Meier survival estimates

Proportion not yet returned

Time since migration to the US

Work Channel — Education Channel
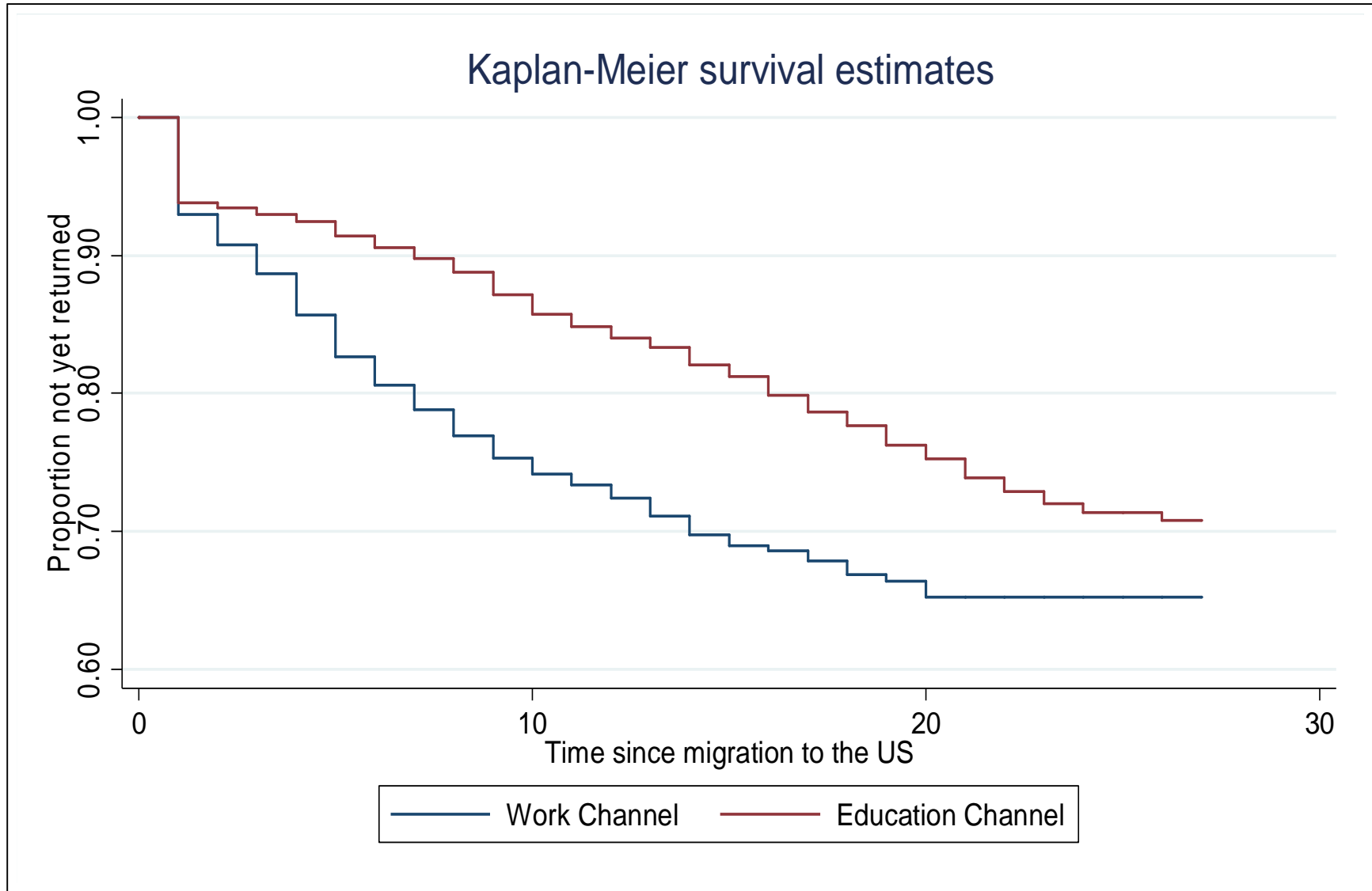
Methodology

- Longitudinal information → Survival analysis

- Allows for time-varying covariates

- Can predict the timing of the return decision, not only prob. of occurrence

    → Is probability of return time-dependent?

- H1: negative time-dependence, evidence of negative skill based self-selection (Contant and Massey, 2002)

Methodology

- Discrete time duration analysis, by means of Cox's proportional hazard functions:
$$h(t,x)_i = h(t)\exp(\beta_i X_i)$$

- Two alternative specifications for the hazard ratio $h(t)$:

  ✓ Fully parametric: $h(t,x)_i = exp(\alpha_1 t + \alpha_2 t^2)\exp(\beta_i X_i)$

  ✓ Semi-parametric: $h(t,x)_i = exp(\alpha_1 t_1 + \cdots + \alpha_N t_N)\exp(\beta_i X_i)$

- Expect different time dependence for education & work

- Only 1990 & 2000 cohort, with more reliable data

Methodology

- Covariates at entry
  - Inventor's age
  - Educational level (Master or more at migration)
  - Migration cohort (Cohort 2000=1)
  - Patenting stock at migration
- Covariates while in the US
  - Migrant's status (Student)
  - Educational attainment → Master in the US
  - Educational attainment → PhD in the US
  - Educational attainment → MBA in the US
  - Productivity → Cumulative # patents US

Descriptive stats

	Education channel					Work channel				
	Obs	*Mean*	*Std. Dev.*	*Min*	*Max*	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Migration cohort	50211	1993.1	4.630	1990	2000	15333	1996.8	4.648	1990	2000
Age at migration	50211	24.32	2.652	18	52	15333	31.87	5.956	18	62
Master or more at migration	50211	0.09	0.283	0	1	15333	0.34	0.473	0	1
Current student status	50211	0.20	0.403	0	1	15333	0.04	0.192	0	1
Master in the US	50211	0.66	0.474	0	1	15333	0.04	0.202	0	1
PhD in the US	50211	0.20	0.400	0	1	15333	0.01	0.097	0	1
MBA in the US	50211	0.08	0.267	0	1	15333	0.04	0.201	0	1
Patents at migration	50211	0.01	0.114	0	5	15333	0.03	0.354	0	12
Cumulative # patents US	50211	3.83	10.64	0	261	15333	4.71	9.07	0	162

Baseline results – Odd ratios

	(1) Education channel	(2) Work channel
Time from migration	0.881***	0.883***
	(0.0201)	(0.0307)
Time from migration^2	1.005***	1.002
	(0.000830)	(0.00195)
Migration cohort = 2000	1.779***	1.423***
	(0.138)	(0.168)
Age at migration	0.872***	0.899***
	(0.00565)	(0.00467)
Master or more at migration	1.623***	1.154
	(0.227)	(0.136)
Current student status	0.595***	0.160***
	(0.0681)	(0.0809)
Master in the US	0.432***	0.724
	(0.0444)	(0.215)
PhD in the US	0.552***	1.259
	(0.0744)	(0.763)
MBA in the US	0.866	0.401**
	(0.148)	(0.169)
Patents at migration	2.525***	1.429***
	(0.358)	(0.0842)
Cumulative # patents US	1.001	1.011**
	(0.00429)	(0.00528)
Observations	50,211	15,333
Times dummies	NO	NO
# unique inventors	3054	1308

Baseline results – Odd ratios

	(3) Education channel	(4) Work channel
Time from migration		
Time from migration^2		
Migration cohort = 2000	1.867***	1.424***
	(0.150)	(0.170)
Age at migration	0.977	0.904***
	(0.0159)	(0.0115)
Master or more at migration	1.180	1.138
	(0.176)	(0.139)
Current student status	0.459***	0.173***
	(0.0908)	(0.0884)
Master in the US	0.568***	0.719
	(0.0709)	(0.216)
PhD in the US	0.585***	1.430
	(0.0805)	(0.835)
MBA in the US	0.711**	0.403**
	(0.124)	(0.171)
Patents at migration	2.320***	1.431***
	(0.301)	(0.0822)
Cumulative # patents US	0.999	1.012**
	(0.00524)	(0.00528)
Observations	50,211	15,094
Times dummies	YES	YES
# unique inventors	3054	1308

Baseline results

Estimated hazard ratios since entry in the US, by education level - Education migrants, 1990 cohort
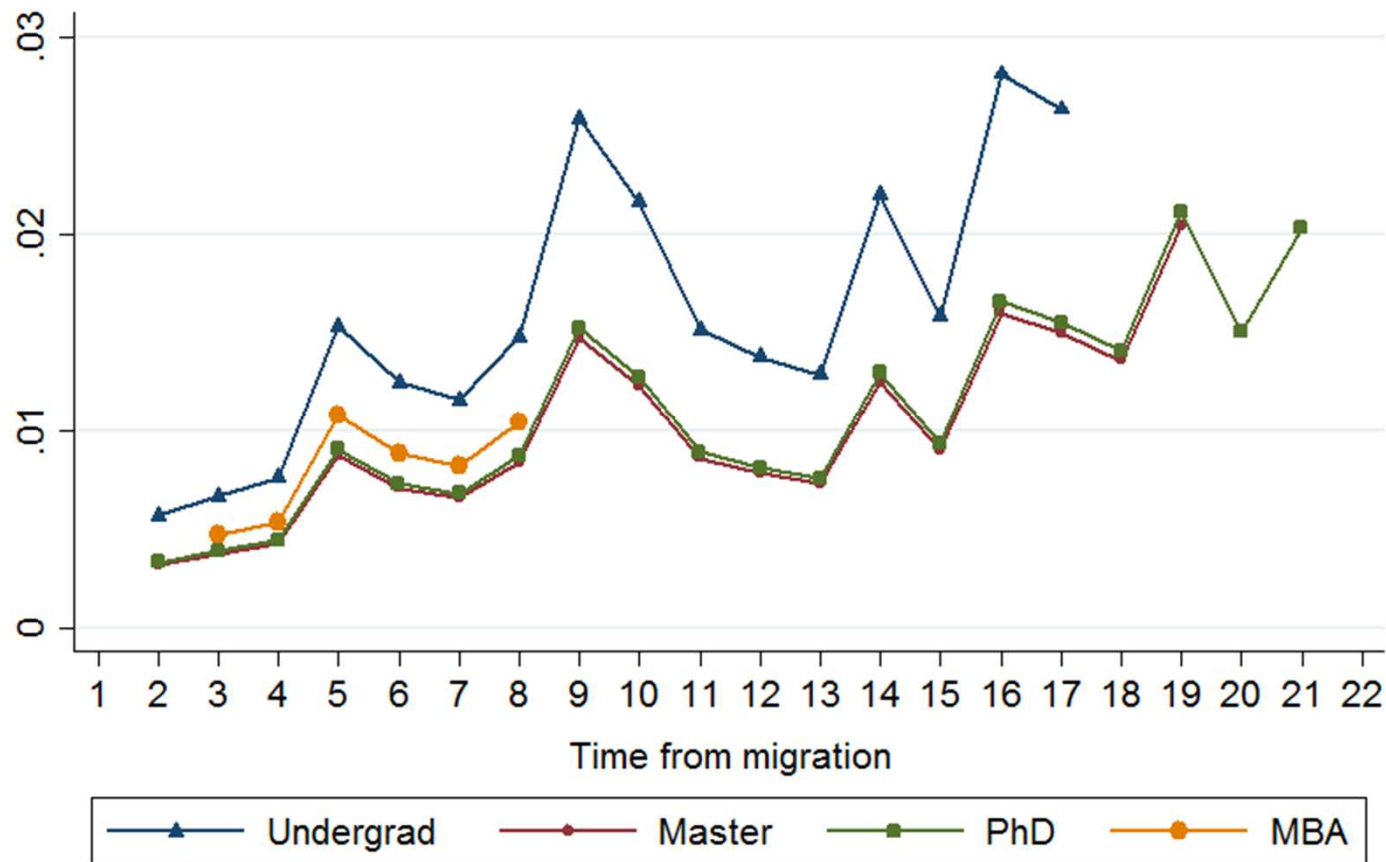
Within sample estimations (unreported regression), for Age at migration =23 and Student status=0 (all remaining regressors at mean values)

Baseline results – Interpretation

- Negative self-selection on education, though not differences between Master and PhD – while only MBA significant for work migrants.

- Positive (and weak) self-selection w.r.t. patenting in the US → different types of inventors: professional vs occasional (specialization) → professionals moving on temporary visas only

- Negative time dependence (for work migrants) → negative self-selection w.r.t. unobservable skills (which are less likely to be applied back in India) → valid interpretation?

- Slight positive time dependence for education migrants→double affiliation?

Concluding remarks

- We've looked at an understudied topic → return migration (data shortage)

- Evidence of education-based negative self-selection conditional on migration channel (education vs work)

- Education migrants, increasingly at risk to return → circular migration? Double affiliation?

- When looking at migrants entering with temporary visas (work/education), the US appears more attractive than Canada or Australia (point-based visa systems) (Koslowski, 2018)

- Temporary visas can be turned permanent → what is the stay rate of highly skilled workers? We show evidence that this is high in the US!!

Concluding remarks

- Since when?

Spotted in Silicon Valley

Further steps

- Are the quality of education institution (both in IN and US) or the labor market and social networks (Silicon Valley effect) important?

- Inter-company vs intra-company moves?

- What about other migration corridors?

- Agrawal's et al. (2011) results (also Breschi et al., 2017)→Indian diaspora does not channel knowledge back: what about returnees?

Return migrants' self-selection: Evidence for Indian inventors

Thank you!

TKC project — http://tkc-anr.eu/

@tkc_anr

Stefano Breschi   **Francesco Lissoni**   **Ernest Miguelez**