Migration of Professionals to the U.S.: Evidence from LinkedIn Data

Bogdan State, Mario Rodriguez, Dirk Helbing, Emilio Zagheni

Stanford University

January 12, 2018

Motivation

- Ongoing debate about immigration in the United States.
- U.S. = not only traditional country of immigration, also top destination for scientists and highly-educated professionals.
- "individuals making exceptional contributions to science and engineering (S&E) in the United States are disproportionately drawn from the foreign born." (Levin & Stephan, 1999)
- Traditionally little attention paid to position of United States in global system of migrations.
- Unique opportunity to study how U.S. is faring in the global competition for talent, as well as general dynamics of highly-skilled migration.

Global Trends in Highly-Skilled Migration

- Since early 1990s, large increase in migration (Anich et al., 2011; UNDP, 2013).

- Global jump in Highly-Skilled Migration (Lowell, 2008).

- Increased friction for HSM to the United States:
  - H1B visa caps.
  - Country-level quotas for permanent residency.
  - (More recently) Interview requirements for permanent residency.

- Expectation: emergence of other destinations for HSM.

- Can we measure it?

New Migration Data Paradigm

- Migration data of fundamental importance to policy.
- Traditional data problems:
  - coarse
  - inconsistent
  - expensive
  - slow
- Geolocated data as complement to traditional sources:
  - Cellphone call records.
  - IP geolocation.
  - Phone-based geolocation.
  - Many other sources!

A unique set of trade-offs

- Digital migration datasets can be evaluated along three axes:
  - scope
  - time-series length
  - geolocation accuracy
- Hard to do well on all 3 criteria!
- LinkedIn Data
  - Near-global scope.
  - 20+ years (unique feature).
  - Country-level accuracy.

LinkedIn Dataset

- Social networking website for professionals.
- At time of study, about 200m monthly active users, currently 500m+.
- LinkedIn profiles contain employment and educational history, often stretching back decades.
- Aggregated, anonymized data could be used to estimate changes in highly-skilled migration.
- Retrospective dataset: period examined covers time prior to creation of LinkedIn platform.

Methodology

- Initial population of 200m LI users worldwide.
- Some employment spells present in LinkedIn data have self-reported geolocated information (i.e. country of employment).
- Problem: insufficient data.
- Used employment spells as training data for Naive Bayes classifier to predict country of employment for remainder positions.
- At decision threshold, 99% precision, 54% recall against held-out dataset.
- Extracted inter-country migration events.
- Also matched job titles against Standard Occupational Classification (SOC) (for STEM status).
- Educational background also available, could even match against Quacquarelli-Symonds top 500.
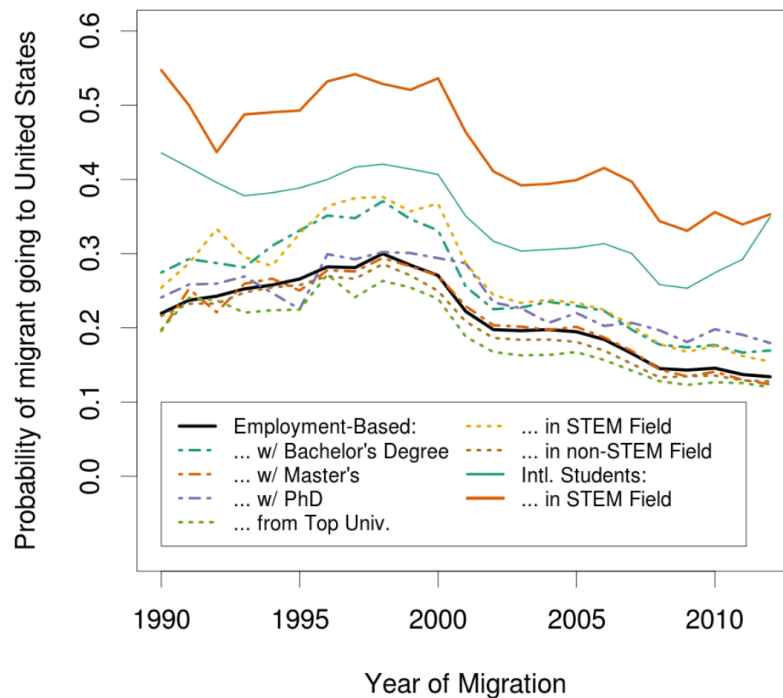
Main Findings

Figure 1: Conditional Probability of Migration to United States by Year, 1990-2012.

Main Findings

- Data shows **conditional probability** of migrating to the United States, as opposed to another country, for HSMs.

- Slight increase during 1990s, subsequent decline.

- Shows that proportionally fewer HSMs go to the U.S. out of all HSMs worldwide.

- Replicated across educational levels, STEM / non-STEM and for Shangai top 500 graduates.
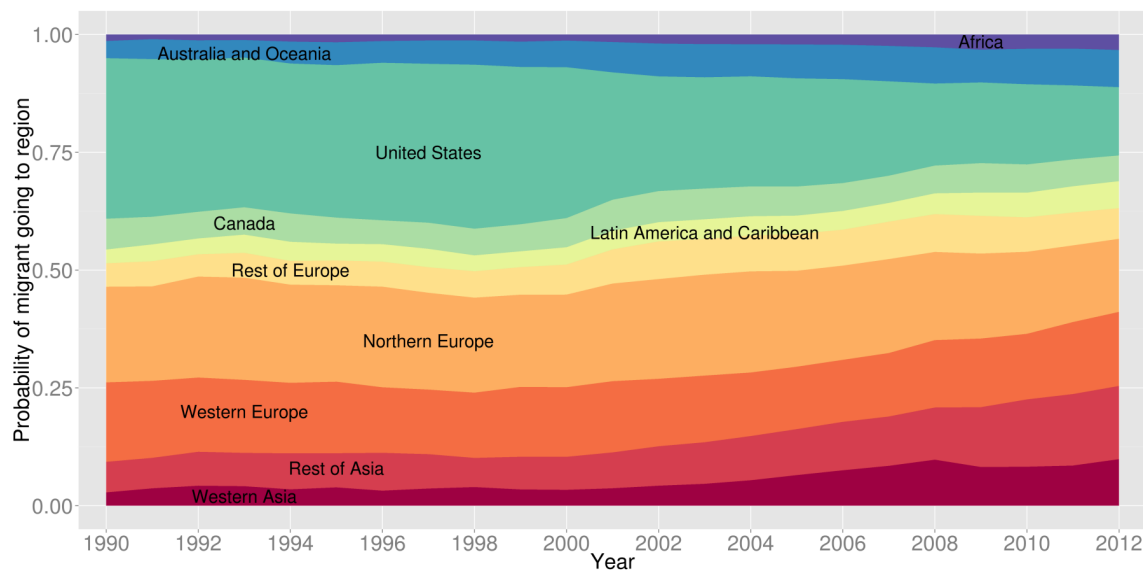
Main Findings

Figure 2: Distribution of Migration Flows, by year and region of destination, 1990-2012.

Main Findings

- As of time of writing: "The United States continues to occupy a central place in the global migration system. However, its dominant position is no longer indisputable."

- Europe and Canada also saw decrease.

- Increase for Australia, Oceania, Africa and LatAm.

- Most prominent increase for Asian countries (cumulative conditional probability of 25% in 2012 vs 10% in 2000).
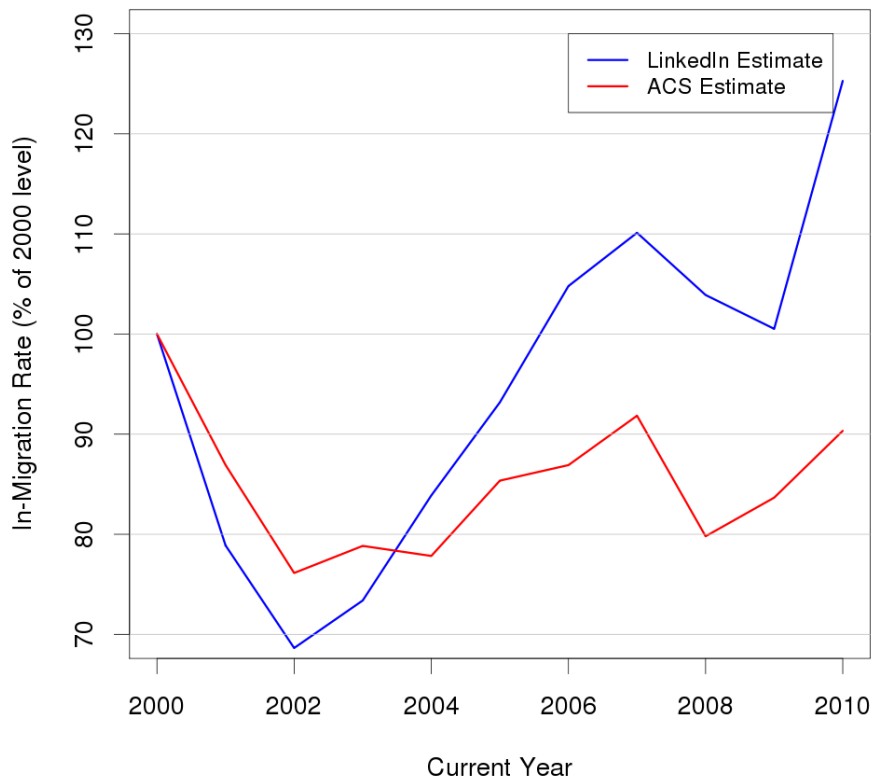
Robustness Checks: Cohorts

- Less concerning limitations:
  - Do not know actual citizenship status.
  - What happens with circular migrations?
- Very concerning limitation – representativity.
- LI users not random sample!
- Alternative explanation: LI is US company, earlier cohorts more likely to be in US.
- Stat-sig. lower post-2000 prob. of move to US replicates for 10 yearly user sign-up cohorts.

Robustness Checks: ACS

- Also possible to compare with American Community Survey data on immigration for individuals with ≥ bachelor's degree.
- highly-correlated for 2000-2005.
- LI indicates higher HSM post 2005.
- Either ACS underestimates professional in-migration (inconsistent treatment of "non-immigrant" visas?) or effect revealed by LI should be even stronger.

Robustness Checks: ACS

Robustness Checks: Top Universities

- Who is "really" a highly-skilled migrant.
- Alternative explanation: dilution of educational quality of individuals with higher-ed. degrees who migrate.
- Possible to control for institutions being canonical list of top 500 universities (i.e. Quacquarelli-Symonds).
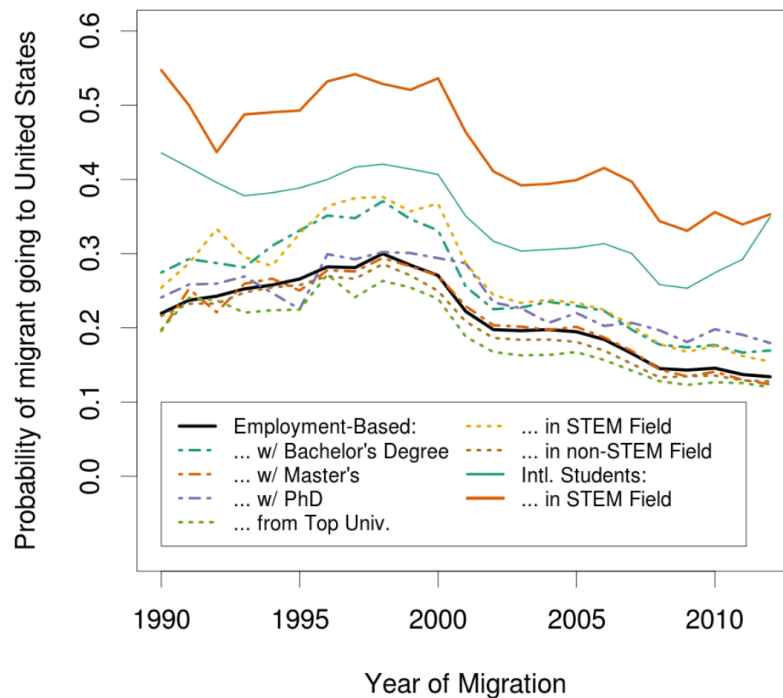- Same trend visible!

Figure 4: Conditional Probability of Migration to United States by Year, 1990-2012.

Conclusions

- Evidence for HSM becoming more diverse w/r to destination countries, with U.S. attracting proportionally fewer.
- May be a reflection of:
  - increased competition for highly skilled migrants from other countries.
  - of declining demand for highly skilled migrants in the U.S.
- Findings robust to several tests for common alternative explanations.

Promise of Digital Data for Highly Skilled Migration

- Traditional data for HSM even sparser than for traditional migration flows.
- Porous concepts of what both "migration" and "highly-skilled" represent.
- More timely and accurate understanding of highly-skilled migration patterns possible with LI data.
- Challenge: how well does this match traditional data?
  - If we had traditional data, why do the digital measurement? (low novelty)
  - If we don't have traditional data, how can you trust the digital measurement? (low reliability)
- Need to strike a balance, combine digital and traditional datasets.

References

Anich R., G. Appave, C. Aghazarm, F. Lacko, and A. Kigouk. *International Migration Trends, World Migration Report Series*, 2011/02, IOM. Chapter 2 (2011).

Levin S.G., and P.E. Stephan. *Are the Foreign Born a Source of Strength for U.S. Science?* Science, 285 (5431), 1213-1214 (1999).

Lowell L. *Highly-Skilled Migration*, World Migration Report Series, 2008/04, IOM (2008). Chapter 2.

Quacquarelli Symonds Ltd., *World University Rankings*, August 1st 2013 http://www.topuniversities.com/qs-world-university-rankings.

United Nations Population Division/DESA. *Trends in International Migrant Stock: The 2013 Revision* (2013).