Chinese Graduate Students and US Scientific Productivity, Evidence from Chemistry

Patrick Gaule and Mario Piacentini

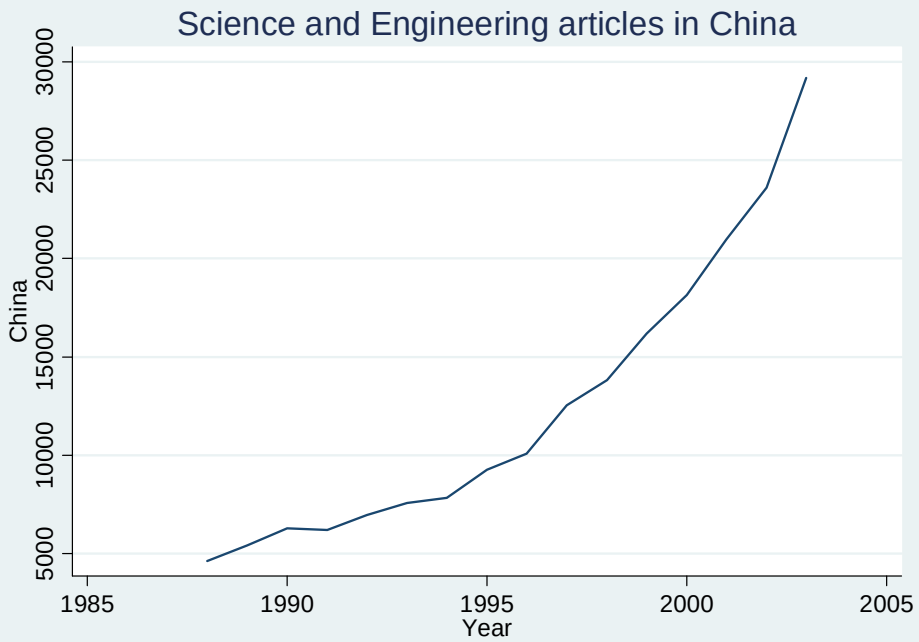
MIT & University of Geneva

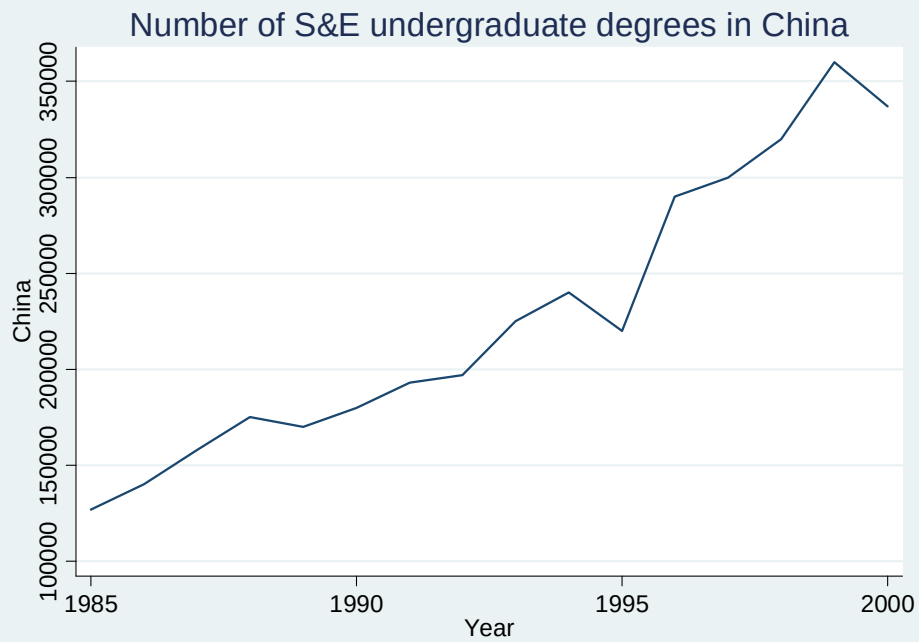
September 2009

Motivation

- The recent flow of Chinese students to US universities is one of the most relevant episodes of skilled migration so far
  - Of 26'891 students graduating with a PhD degree in S&E in the United States in 2003, more than 2500 (9.2%) were born in China (NSF 2007)
  - PhD students graduating in 2006 from US universities were more likely to have done their undergraduate studies at Tsinghua University or Peking University than at the University of California, Berkeley, or any other institution (Melvis 2008)
  - As of 2000, 8.9% of doctorate holders in U.S. science and engineering occupations were born in China (NSF 2007)
- Chinese students have the highest rate of stay after the end of their PhD program, thus representing "stable" inputs to the US stock of human capital

China's Rising Strength





Is the immigration of Chinese students a good thing for the US?

- One might expect a massive inflow of highly talented individuals to be beneficial for the United States (Freeman 2009)
- Some counter-arguments:
  - Chinese and other immigrants may crowd out US-born students from graduate programs, either because there is a limited number of spots available or by depressing wages of PhD-educated scientists and engineers (Borjas 2004, 2006)
  - Graduate students typically do not bear the full cost of their training, so that the US are effectively subsidizing the tertiary education of foreign-born students (Borjas 2004)

- Literature showing the importance of the contribution of foreign-born to US science: Levin & Stephan (1999), Kerr (2008b), Black and Stephan (forthcoming).

- Hunt (2009) uses the National Survey of College Graduates to show that immigrants publish more scientific articles

- Stuen, Mobarak & Maskus (2007) use a large dataset of publications over 1968-2003 and instruments for student enrollment. They find an effect of the number of doctoral students on productivity at the departmental level but no differential between US and foreign students.

Our contribution

- Regression on productivity at the individual level
- (In progress) Investigation of the role of co-ethnicity with advisor as a determinant of productivity of PhD students.
- Innovations in terms of data
- Extensive checks of quality of ethnic identification and productivity measures

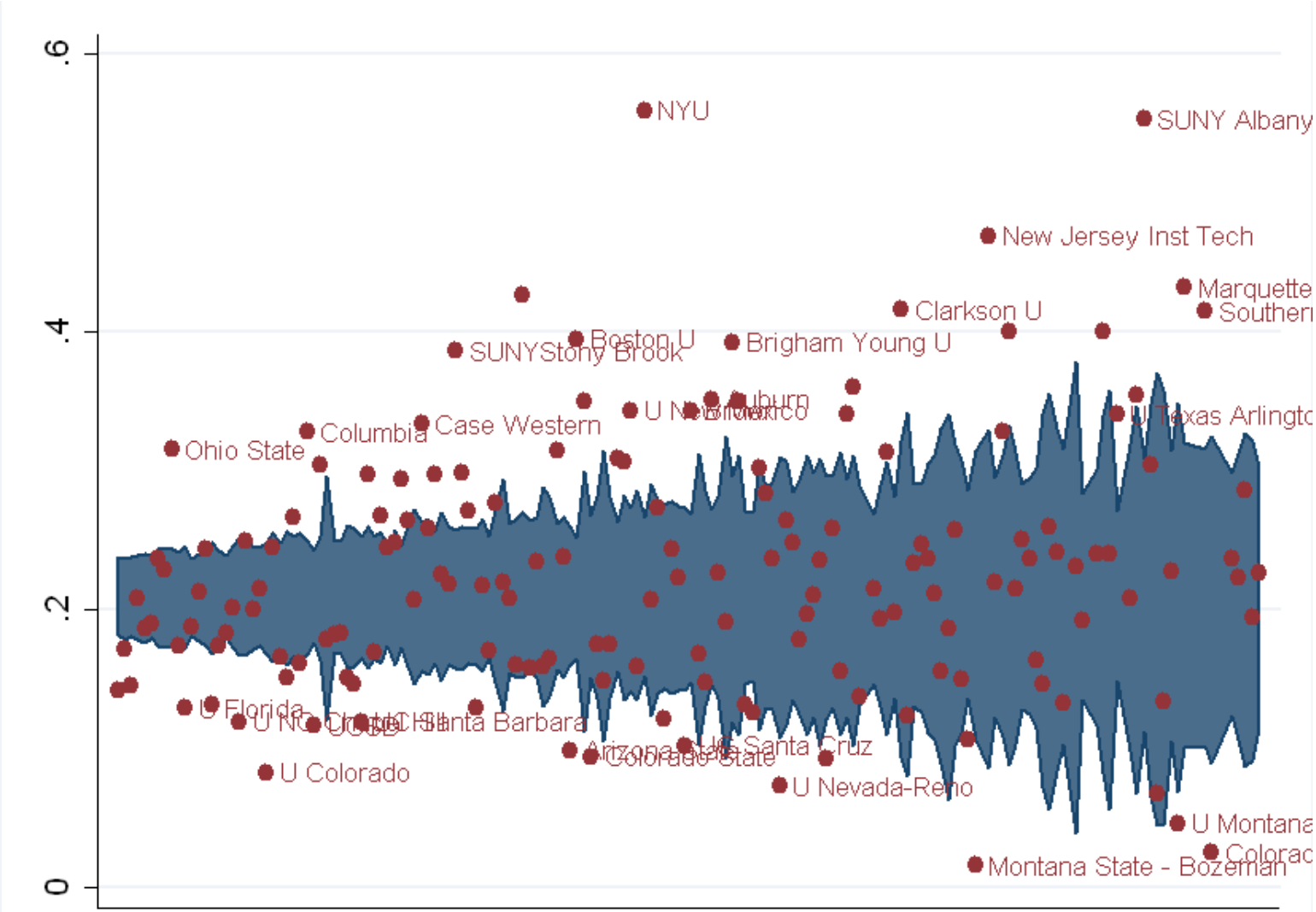
Data (1)

- We use Proquest Dissertations and Theses, to identify PhD graduates
  - Pro: Public database, extensive coverage
  - Cons: No information besides title, name, university, graduation year and advisor (except for some universities)
- We match the list of PhD students with a list of common Chinese last and first names to identify Chinese students (technique introduced by Kerr (2008)).
- Match with publication data from Scopus to construct individual productivity measures. .

Data (2)

- We have about 32k PhD graduates in Chemistry from US universities between 1999 and 2008 (more than 95% of the population of graduates)
- About 13,7% of graduates have Chinese first and last names. This fraction does not vary much across graduation years
- Uneven distribution across universities

Sorting across universities



Quality of matching

We use a subsample of our data with biographic information to check the quality of ethnic identification through family and first names.

Table 1 - Quality of name matching

	With Chinese name
Right Match	90.21% are educated in China
Wrong Match	9.79% not educated in China (5.23 Taiwan, 3.49 US)

	Educated in China
Right Match	91.19% have a Chinese name
Wrong Match	8.81% have not a Chinese name

Table 2 - Descriptive Statistics

	Chinese (n=4,404)		non-Chinese (n=27,656)	
Variable:	μ	σ	μ	σ
Any pub	0.280	0.449	0.232	0.422
# of Pub	0.519	1.092	0.378	0.862

Note: First authored publications only

	Number Pub	Number Pub	Number Pub
Chinese name	0.130***	0.0762***	0.120***
	[-0.015]	[-0.020]	[-0.015]
Late		0.264***	
		[-0.0231]	
Late*Chinese		0.113***	
		[-0.029]	
Advisor			0.162**
			[-0.0659]
Advisor*Chinese			0.225**
			[-0.104]
Constant	0.155***	0.159***	0.155***
	[-0.0376]	[-0.0376]	[-0.0376]
Controls	yes	yes	yes
Year, Field and Uni fixed eff	yes	yes	yes
Observations	29034	29034	29034
R-squared	0.036	0.036	0.037

Notes: Robust standard errors in brackets.*** significant at 1%

First or second generation? Results on biographic data

	(I) Number Pub	(II) Number Pub
Chinese name	0.187***	
	-0.0339	
Undergraduate in China		0.179***
		-0.0344
Constant	0.221**	0.223**
	-0.0967	-0.0967
Graduation year fixed effects	yes	yes
Controls	yes	yes
Subfield fixed effects	yes	yes
University fixed effects	yes	yes
Observations	6091	6091
R-squared	0.049	0.048

Notes: Robust standard errors in brackets.*** significant at 1%

Summary of the results

- PhD students having a Chinese last and first name publish more during their studies, controlling for the characteristics of the programs they are enrolled in.
- The productivity gap is higher in recent years. This might be explained by:
  - an increase in the relative quality of the Chinese students
  - higher motivations due to Visa reasons (decreased proportion of permanent visas)
- Preliminary evidence shows that Chinese students benefit more from having a Chinese PhD supervisor.
- Results hold in a restricted sample for which Chinese are identified as those having studied in China (thus excluding second generation immigrants).

Robustness checks

- Results hold when considering other publication-based measures of productivity.
  - Notably, Chinese have also an higher number of **non-first** authored publications.
  - The publications of first authors with a Chinese name receive more citations
- The tendency of Chinese to have frequent last names might bias the match between student and publication data. However, we show that homonymity does not drive our results.
- Considering US PhD-granting chemistry departments, 23% of first authors have a Chinese last name. However, only 21.8% of graduate students, 14% of post-doc (roughly) and 6% of professors have Chinese last names.

Conclusion and discussion

- Chinese graduate students in US chemistry departments have a higher productivity than their colleagues

- Why?
  - Higher effort possibly induced by higher opportunity cost of not succeeding. Survey evidence showing that Chinese post-doc spend more time in the lab than American post-docs (Sigma Xi survey).
  - More stringent selection process, either in China and to get into the US.

- Policy implications
  - Our finding strengthens the case that the US are benefiting from the influx of Chinese students.
  - The demand from Chinese for US graduate education will continue to increase as access to higher education in China is expanding. US Universities could benefit from strengthened ties with Chinese higher education institutions.

THANK YOU

Addressing homonymity

- Matching on last name, first initial, university and department
- Excluding known duplicates (individuals with common 4 identifiers) among graduate students and professors
- Robustness check: dividing the sample into frequent last name/first initial combinations and less frequent combinations