

# Chinese graduate students and US scientific productivity: evidence from chemistry\*

Patrick Gaulé<sup>†</sup> & Mario Piacentini<sup>‡</sup>

January 17, 2010

## Abstract

Large numbers of Chinese-born students are graduating from US universities with PhD degrees in Science and Engineering. We show that Chinese graduate students are more productive than other graduate students in the same department and university. The productivity differential is higher for more recent years and for Chinese students working with Chinese advisors. We relate the higher productivity of Chinese students to competitive selection both in China and to get into US PhD programs. Our results strengthen the case that the US is benefitting from the influx of Chinese graduate students.

*DRAFT- Please do not quote or distribute*

---

\*We thank Ajay Agrawal, Pierre Azoulay, Michel Cincera, Dominique Foray, Bill Kerr, Jacques Mairesse, Megan McGarvie, Fiona Murray, Maurice Schiff, Paula Stephan, Scott Stern, Yanbo Wang and seminar participants at the NBER productivity lunch, the COST workshop on the economics of science and the second international conference on migration and development (10-11 September 2009, Washington DC) for insightful comments, discussions and advice. Maria Eugenia Sagon provided excellent research assistance.

<sup>†</sup>Sloan School of Management, Massachusetts Institute of Technology, pgaule@mit.edu

<sup>‡</sup>Department of Economics, University of Geneva, mario.piacentini@unige.ch

*“The conquest of the technological frontier, like the conquest of the geographical frontier, requires [...] millions of individuals”*

(Milton Friedman, 1955)

## 1 Introduction

Immigrants from China are a large, and increasing, fraction of Science and Engineering PhD graduates educated in the United States<sup>1</sup>. Of around 27'000 students graduating with a PhD degree in Science and Engineering 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 (Mervis 2008). 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 and to the US Science and Engineering workforce<sup>2</sup>. As of 2000, 8.9% of doctorate holders in US Science and Engineering occupations were born in China (NSF 07).

These numbers reflect a fundamental global imbalance in higher education. The supply of PhD-level Science and Engineering education in research-intensive programs is largely dominated by US Universities. However, foreign countries account for a very large, and increasing, fraction of the demand for such education. In 2007, close to 1.7 million scientists and engineers obtained undergraduate degrees from Chinese universities compared to less than 0.5 million from US universities. The United States has a liberal immigration policy for students and US universities are enrolling many foreign-born graduate students. Nevertheless, it remains considerably harder for foreign-born undergraduate students to access high quality graduate education than it is for their American counterparts.

The welfare effects of the migration of young scientists and engineers are not well understood<sup>3</sup>. The influx of Chinese graduate students may be beneficial for the United States by reinforcing the US comparative advantage in high technology and university workforce intensive sectors (Freeman 2009) or by increasing the global rate of knowledge production. However, a number of counter-arguments have been raised. First, 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). Second, graduate students typically do not bear the full cost of their training, so that the US is effectively subsidizing the

---

<sup>1</sup>Since the passing of the Chinese Exclusion act in 1882 and until the late 1960s, the numbers of Chinese immigrants to the US were minimal. Until 1978, Chinese migration policy only allowed migration into socialist countries. The enactment of the Immigration and Nationality act in 1965 gave start to a period of steadily rising Chinese migration to the US. In 2002, of the more than 700 thousands temporary immigrants to the US from China, about one-tenth of them are students (USCIS data reported in Poston & Luo (2007))

<sup>2</sup>According to estimates derived by Finn (2007) using Social Security data, the stay rate for Chinese doctorate recipients is around 92 percent after five years from the PhD, the highest observed for any country in 2005.

<sup>3</sup>Only destination countries effects are discussed here. Source country effects are equally important and interesting.

tertiary education of foreign-born students (Borjas 2004). These counter-arguments could become particularly relevant if return migration becomes more frequent in the future.

The main contribution of the paper is to show that Chinese students graduating from US chemistry PhD programs write more and higher quality papers than other students during their thesis. This evidence is highly relevant for the assessment of the benefits for destination countries of admitting large number of foreign young with student visas. While it is already clear that the large demand for graduate positions in the US has helped universities to get staffed at lower costs, this new evidence on productivity can strengthen the arguments to keep the doors of the foreign student programs open. Moreover, there is a large heterogeneity in the quality of students coming from different countries, and the increasing primacy of Chinese students is clearly linked to the rising quality of training and of student selection in China.

We use individual level data on the whole population of students completing a PhD program in over 200 top US chemistry departments. We identify the ethnic origin of these students by matching the data with an extensive list of Chinese family and first names. Productivity is measured by the number of first-authored papers published during the PhD weighted by the impact factor of the publishing journal(s). There is strong evidence that the Chinese students have a higher productivity, controlling for fixed unobserved and time varying observed characteristics of the programs they are enrolled in. The productivity differential is higher for more recent years and for Chinese students working with Chinese advisors.

Two main explanations of the finding on the productivity of graduate students can be advanced. Chinese students who manage to obtain a spot in a US PhD program are selected among a large pool of potential applicants who have themselves passed through the highly competitive Chinese education system. The second explanation has to do with career opportunities. It could be that pursuing an academic career is relatively more attractive for Chinese students, either due to immigration rules or to discrimination in non-academic occupations.

Our findings are related to a recent and growing literature on the role of foreign-born scientists and engineers in US Science. Levin & Stephan (1999) show that individuals making exceptional contributions to Science and Engineering in the United States are disproportionately drawn from the foreign born. Hunt (2009) uses the National Survey of College Graduates and finds that immigrants publish more scientific articles. Kerr (2008b) matches inventors of USPTO patents residing in the US with a ethnic name database and finds an important share of inventors of Chinese and Indian ethnicity (13.9% over the period 2000-2004). Black and Stephan (forthcoming) use the same ethnic-name matching techniques for authors (based in the US) of *Science* papers and similar find an important share of papers signed by authors with non-English and non-European names.

Closest to our paper is the study of Stuen, Mobarak & Maskus (2007). Using 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.

The rest of the paper is organized as follows. In section II we describe our data and present descriptive statistics on enrolment and productivity of Chinese graduate students. Results are in section III. Section IV concludes.

## 2 Data

Our dataset is based on a list of PhD graduates in chemistry and chemical engineering derived from the Proquest Dissertations and Theses database. This database lists the name, title, graduation year, university and advisor of PhD graduates from major US universities<sup>4</sup>. Considering the graduation years 1999-2008, we have a total of 32,069 individuals across all fields of chemistry (organic chemistry, biochemistry, analytical chemistry, chemical engineering, inorganic chemistry, nuclear chemistry, geochemistry).

*Identifying Chinese students.* An important feature of our dataset is that we do not have systematic information on citizenship or place of birth for PhD graduates. Instead, we identify Chinese students through a list of common Chinese last and first names. This approach of ethnic name matching, introduced by William Kerr (2008a) in his study of co-ethnicity in patent citations, has distinct disadvantages. Specifically we cannot perfectly distinguish between first- and second generation immigrants or between students originating from mainland China and Taiwan. On the other hand, ethnic name matching enables us to analyze large data sets from publicly available sources.

We are concerned about incurring two types of errors when identifying ethnicity through name lists. The first type of error concerns those students who have a Chinese name according to our lists of first and last names, but were not actually born in China. This is most likely due to the commonalities between Chinese names and names from Taiwan, Singapore, Hong Kong and Korea. Distinguishing between the first and subsequent generation of migrants is also an issue. The second type of errors concerns those students who are born in China, but fail to be identified as Chinese as they have a name or first name that is too rare to be included in the list. The Chinese migrants who change their first name with an occidental one are also at the origin of this second type of error. We use a subsample of universities, covering about 20% of the sample where we have detailed biographic information for all PhD candidates<sup>5</sup> to check to what extent our ethnic identification technique suffers from the two types of errors. It turns out around 90% of those identified as Chinese from our name list had done their undergraduate studies in China (a further 5% had studied in

---

<sup>4</sup>All universities with very high research activity according to the Canergie classification- except Yeshiva university- and around 3/4th of universities with high research activity.

<sup>5</sup>Biographical information in theses is used in MacGarvie (2007)

Taiwan). The second type of error is approximately of the same size: around 91% of those educated in China are identified as Chinese through our name matching technique.

(table 2 about here)

*Allocation of Chinese students.* Out of 32,069 PhD graduates in chemistry, 4,413 or 13.76% have a Chinese name. This fraction is more or less constant across graduation years. However there is considerable variation across universities. Lower tier schools (based upon R&D expenditures) tend to have relatively more Chinese students.

(figure 1 about here)

To check whether a random allocation could have generated the observed dispersion of Chinese graduate students, we ran one thousand simulations whereby individuals were randomly assigned to universities taking into account university size<sup>6</sup>. From the simulated data, we can calculate a 95% confidence interval on the proportion of Chinese graduate students at each university<sup>7</sup>. We report these confidence intervals together with the observed allocations in figure 2.

(figure 2 about here)

Figure 2 provides evidence of ethnic sorting of Chinese across universities. We are currently investigating factors that might explain such ethnic sorting, with preliminary results suggesting a role for the proportion of ethnic faculty (faculty with a Chinese last name) and for the extent to which universities provide financial support to their graduates students.

*Constructing productivity measures.* To construct individual productivity measures, we match our list of PhD graduates with publication data from Scopus. Our preferred measure of productivity is the the number of first-authored papers published during the PhD<sup>8</sup> weighted by the impact factor of the publishing journal(s). In chemistry, first-authorship is the most prestigious spot<sup>9</sup> and is usually given to the individual who has conducted most of the experiments. Alphabetical ordering of authors is not a common practice- we checked with our data that individuals whose name starts with an 'A' were not more likely to be first-authors than those with a name starting with a 'Z'. Table 3 displays different measures of productivity for Chinese and other students: whether the graduate

---

<sup>6</sup>In each simulation, we had 32'000 graduate students of which 4'500 Chinese and 27'500 non-Chinese randomly allocated to universities according to their For instance, Stanford had 542 PhD graduates in chemistry during this period, or 1.54% of the total number of PhD graduates in chemistry. So in the simulation each of the 35'000 individuals had a 1.54% chance of being allocated to Stanford. For a specific simulation, we calculate the proportion of Chinese at each university

<sup>7</sup>the interval is constructed as the mean plus and minus 1.96 standard deviations of the simulated data so as to include 950 of the 1000 simulations

<sup>8</sup>We take publications form three years before the graduation year to one year after the graduation year.

<sup>9</sup>In his book 'The Road to Stockholm', Istvan Hargittai reports (2002:231) that Martin Kamen had 'generously yield first authorship to Ruben since he badly needed it for promotion' and this could have been a factor in Kamen not receiving the Nobel prize for his discovery of carbon-14.

student has published any paper as first-author during his PhD, the number of papers first-authored and the same weighted by the 2008 impact factor of the publishing journal. To address concerns of homonymity, we exclude from the sample graduate students who share the last name and first initial with another graduate student, or a professor, of his university and department<sup>10</sup>.

(table 3 about here)

Table 3 presents descriptive statistics on the scientific productivity of graduate students, distinguishing Chinese students and other students. The raw numbers suggest a productivity advantage for Chinese graduate students which we study in more detail in the next section.

### 3 Estimation and results

We present our results using the number of first authored publications as productivity measure. We regress it on a dummy that take the value one for graduate students with a Chinese last and first name. We use university fixed effects to control for heterogeneity at the university-level and in particular differences in quality across universities. Subfield and year of graduation fixed effects are also included. All estimates are ordinary least squares estimates with robust standard errors. The results are qualitatively the same using other measures of academic productivity (whether the graduate student has published any paper as first-author during his PhD, and the number of papers first-authored weighted by the 2008 impact factor of the publishing journal).

(table 4 about here)

In all three regressions reported in columns I, II and III, we find that graduate students with a Chinese last name are more productive and the effects are significant at the 1% confidence level. We postpone the interpretation of this effect to the next section.

In column II, we examine whether the productivity advantage of Chinese has increased over time, adding the dummy late which refers to the graduation years after 2003. The productivity of all PhD students has increased, but those of the Chinese relatively more, as evidenced by the interaction term.

We also introduce tentative evidence on whether Chinese students are more productive when collaborating with faculty of Chinese origin. We collected data linking students to advisors to address this point directly. Around forty percent of the students of Chinese advisors are themselves Chinese. As can be seen in column III, Chinese students benefit relatively more from the supervision of a Chinese professor<sup>11</sup> and the effect is surprisingly large. Two explanations can be advanced.

---

<sup>10</sup>This results in a loss of less than 3% of observations. Homonymity concerns are further discussed in the appendix.

<sup>11</sup>Chinese Professors are identified as those born or having done undergraduate studies in China. This is evidenced by the interaction term. A Chow test confirms the difference of the effect of a Chinese advisor for Chinese and non Chinese student (the test yields a value of 3.32, rejecting equality of coefficients at 10% level of significance).

First, Chinese advisors may be either attracting or selecting particularly talented Chinese students. This could happen for instance through better networks and better knowledge of universities at origin. The second is that coethnicity between advisor and student may lower communication and assimilation costs of Chinese students.

We already stated that one problem of our ethnic identification technique is the difficulty of distinguishing first from second generation migrants. We verify whether this issue is relevant for our results using the subsample of the data for which we have biographic information.

(table 5 about here)

As shown by Table 5, we get similar results if we either consider Chinese students as identified by our name matching technique, or Chinese students identified as those who did their undergraduate studies in China, thus excluding second generation migrants. Moreover, the effect is larger if we consider those educated in China than those with a Chinese name and when both are included in the same regression only undergraduate education in China is significant.

## 4 Discussion and conclusion

The central result of this paper is that Chinese graduate students have on average a higher productivity during their thesis than other graduate students in US universities. Our preferred explanation for this result is a selection effect. US education enjoys an excellent reputation in China and attracts the brightest and most motivated Chinese students. Despite the fact that US universities are admitting large number of Chinese students, it is nevertheless considerably more difficult for a young Chinese scientist or engineer to obtain admission into a US PhD program. Evidence from Attieh and Attieh (1997) suggests that top US universities give substantial preference to US citizens in their admission processes.

Having graduated from a top Chinese university is a *de facto* requirement for entering a US PhD program. Indeed Chinese graduate students overwhelming come from a very restricted set of Chinese universities which are extraordinarily selective. Around 10 million high school finishers take the national college entrance exam but only three thousand are admitted into the two most prestigious schools, Peking University and Tsinghua University. Peking University and Tsinghua University are thus more selective than the most electives US institutions- the majority of MIT undergraduates would not have had standardized test scores high enough to be admitted into the undergraduate programs of Peking University and Tsinghua University<sup>12</sup>.

A different explanation for the productivity effect is that Chinese students have a different set of career options than other students. A career in the academia and a post-doctoral training in

---

<sup>12</sup>The median maths SAT score of MIT undergraduates is 770 which is lower than the 1st centile cutoff. Only 3% of Chinese entrance test takers scoring in the top centile are admitted into Peking University and Tsinghua University.

particular may be relatively more attractive to Chinese students (for instance due to immigration considerations), thus increasing the incentives to publish during the PhD. Stephan and Ma (2005) find that temporary visa holders were more likely to pursue postdoctoral training, an effect that was even more pronounced for the Chinese (Stephan personal communication 2009).

Finally, there is anecdotal evidence that Chinese graduate students spend more time in the laboratory. However a survey of post-docs found only small differences between Chinese and Americans in terms of hours worked (50.5 hours per week versus 49.8; Brumfiel 2005).

If one believes that the productivity advantage of Chinese students is mainly due to selection of particularly talented individuals, it is tempting to relate the increased productivity gap over time to larger and better educated later cohorts of Chinese students demanding admission in US PhD programs. Both the investment in higher education and the number of people enrolled has greatly increased in China in the last ten year (Li et al, 2008). If one instead believes that the productivity advantage of Chinese has mainly to do with career options, it is interesting to note that the proportion of Chinese students with a permanent visa has been decreasing<sup>13</sup>. Thus the pressure to publish to ensure a possibility to stay in the US through a post-doctoral training may be increasing.

At a very micro level, our results imply that US chemistry departments that aspire to maximize their production of science should at the margin increase their intake of Chinese students. More important, however, is the relevance of our results to the debate on whether the US is benefiting from the influx of Chinese graduate students. To the extent that universities are selecting the best candidates for the PhD, US universities would have to accept lower quality applications in the absence of Chinese immigration. Our results go further than this simple general equilibrium argument by showing that the Chinese graduate students that are studying in the US have a higher productivity than other graduate students.

---

<sup>13</sup>A substantial number of Chinese students were granted a permanent visa in 1992, through the Chinese Student Protection Act as a response to the Tian'an Men Square Incident in 1989.

## References

Attieh G & Attieh H (1997) "Testing for bias in graduate school admission" *Journal of Human Resources* 32(3)524-548

Black G & Stephan P (forthcoming) "The Economics of University Science and the Role of Foreign Graduate Students and Postdoctoral Scholars in Clotfelter (ed)" *American Universities in a Global Market* NBER

Borjas G (2006) "Native Internal Migration and The Labor Market Impact Of Immigration" *Journal of Human Resources* 41(2)221-258

Borjas G (2004) "Do Foreign Students Crowd Out Native Students from Graduate Programs?" NBER Working Paper No. 10349

Bound J, Turner S, Walsh P (2009) "Internationalization of U.S. Doctorate Education" NBER Working Paper No. 14792

NSF (2007) "Asia's Rising Science and Technology Strength: Comparative Indicators for Asia, the European Union and the United States" National Science Foundation, Division of Science Resources Statistics. NSF-07-319. Arlington, VA.

NSF (2009) "Science and Engineering Doctorate Awards: 2006" National Science Foundation, Division of Science Resources Statistics. Arlington, VA.

Finn, M G (2007) "Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2005", Oak Ridge, TN: Oak Ridge Institute for Science and Education, 2007.

Freeman R (2009) "What Does Global Expansion of Higher Education Mean for the US?" NBER Working Paper No. 14962

Kerr W (2008a) "Ethnic Scientific Communities and International Technology Diffusion" *The Review of Economics and Statistics* 90(3):518-537

Kerr W (2008b) "The Ethnic Composition of US Inventors" HBS Working Paper 08-006

Hunt J (2009) "Which Immigrants are Most Innovative and Entrepreneurial? Distinctions by Entry Visa" NBER Working Paper No. 14920

Hargittai I (2002) "The Road to Stockholm" Oxford University Press: New York

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

Li Y & Whalley J Zhang S & Zhao X, 2008. "The Higher Educational Transformation of China and Its Global Implications," NBER Working Papers 13849, National Bureau of Economic Research, Inc.

MacGarvie M (2007) "Using Published Dissertations to Identify Graduates' Countries of Origin." Unpublished manuscript prepared for presentation at the NBER Conference on Career Patterns of Foreign-born Scientists and Engineers, November 7, 2007.

Mervis J (2008) "US Graduate Training: Top Ph.D. Feeder Schools Are Now Chinese" *Science* 321(5886):185

NORC (2008) "Baccalaureate - Origins of U.S. Research Doctorate Recipients: 1997-2006"

Poston L & Luo H (2007) "Chinese student and labor migration to the United States : Trends and policies since the 1980s" *Asian and Pacific migration journal* 16(3)323-355

Stephan P & Ma J (2005) "The Increased Frequency and Duration of the Postdoctorate Career Stage" *American Economic Review Papers and Proceedings* 95(2):71-75

Stuen E, Mobarak A & Maskus K (2007) "Foreign PhD Students and Knowledge Creation at U.S. Universities: Evidence from Enrollment Fluctuations" University of Colorado at Boulder

## Appendix: addressing homonymity

Because Chinese graduates students tend to have frequent last names, there is a concern that their greater productivity may in fact come from improper attribution of articles. In any large-scale matching between a list of scientists and a list of publications, mistakes arise because there is not enough information in the publication to be certain of the identity of the authors. The problem is more acute for common last names, hence the importance of the homonymity concern.

We address this both by using more information than is usually available for the match and by finding complementary evidence that is immune to homonymity concerns. On the first point, we match publications and people, not only by last name and first initial but also by university and department. Moreover we exclude from the sample graduate students that have the same last name, first initial, university and department as well as graduates students that have the same last name, first initial, university and department as a professor. Reducing the size of the sample in this leads to a loss of less than 3% of observations. The only remaining concern is that we do not have a list of post-docs. Thus, we may mistakenly attribute publications first-authored by post-docs to graduate students with the same last name, first initial, university and department.

While we do not believe that such mistakes are likely to affect the significance of the productivity results, we nevertheless conduct a number of robustness checks. In particular, we divide the sample into frequent last names/first initials combinations and less frequent combinations. To define frequent last names/first initials combinations, we consider the whole population of chemistry graduate students. It turns out that the coefficient on the Chinese dummy is not larger for the frequent combinations subsample than for the less frequent combinations subsample.

In a completely different approach that is not subject to homonymity concerns, we consider all publications from chemistry departments of US universities in our sample and match the names of the first authors with our list of common Chinese names. 23% of papers are signed by a first author who has a Chinese last name. However, only 21.7% of graduate students, 14% of post-docs<sup>14</sup> and around 6% of faculty have a Chinese last name. In other words, scientists with a Chinese name are over-represented in publications first-authorships which is consistent with the higher productivity of scientists with a Chinese last name. Furthermore, the publications of first authors with a Chinese last name are published in journals with higher impact factor, which also holds in a regression with university fixed effects and year fixed effects. The publications of first authors with a Chinese last name also receive more citations in a regression with university fixed effects, year fixed effects and publishing journal fixed effects.

---

<sup>14</sup>Percentage based on Chinese citizens across all fields (Mervis, 2008) .

## Tables and figures

**Figure 1: Fraction of PhD graduates in chemistry with a Chinese last name by university**



Figure 1: Graph generated from our dataset. Dots represent the observed fraction of graduates with a Chinese last and first name. Universities are ranked by the size of their chemistry department (measured in terms of R&D expenditures)

Figure 2: Sorting of Chinese students across US universities with very high research activity

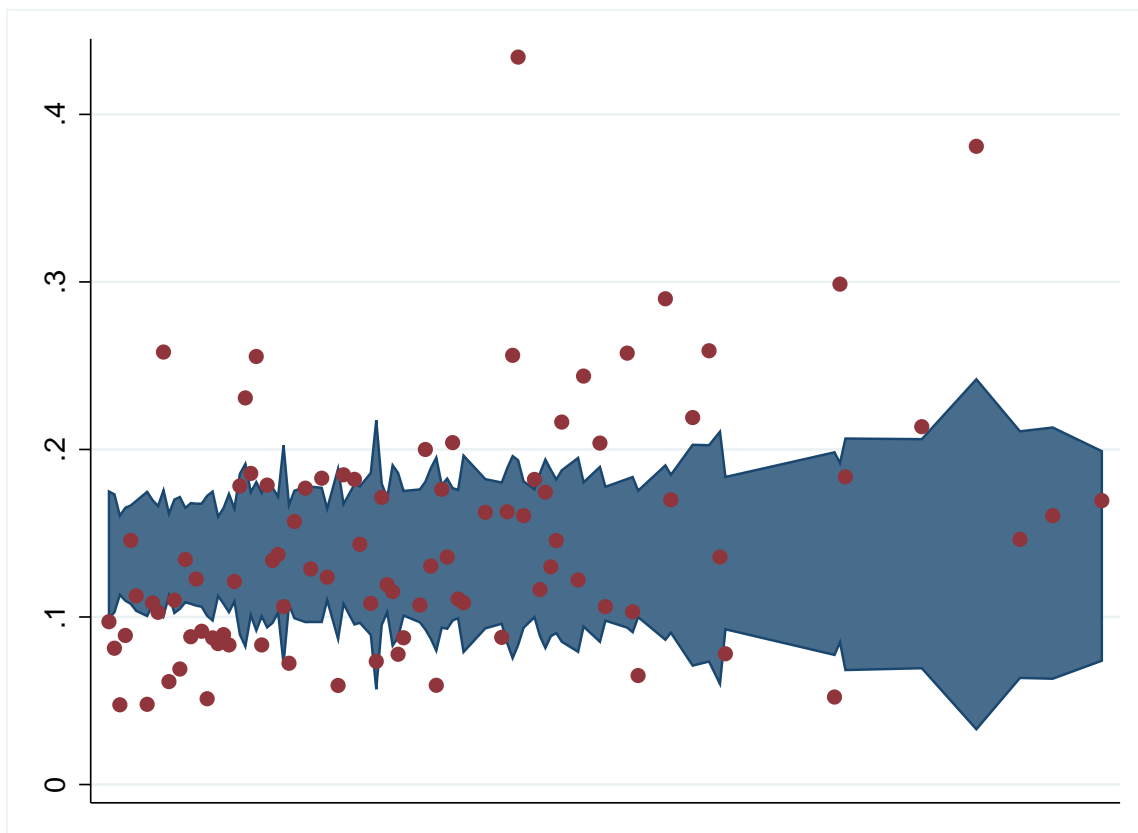


Figure 2: Graph generated from our dataset. Dots represent the observed fraction of graduates with a Chinese last and first name. The area represents a 95% confidence interval derived from 1000 thousand simulations with random allocation of graduate students to universities according to the observed size of their program. Universities are ranked by the size of their chemistry department (measured in terms of R&D expenditures)

Figure 3: Graduating cohorts of scientists and engineers

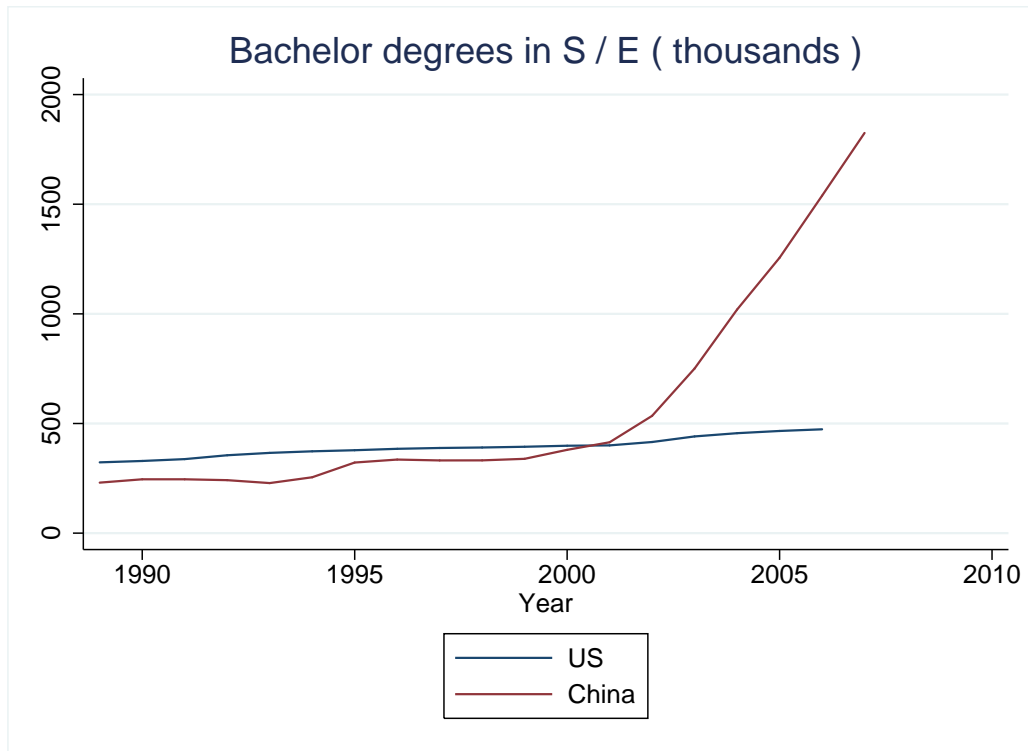


Figure 3: Source: NSF and China Statistical Yearbook, various years

Table 1: Baccalaureate-origin institutions for PhD graduates in S/E from US Universities

Baccalaureate-origin Institution	Country	Rank	All S/E doctorates	Life sciences	Physical Sciences	Engineering
Tsinghua Univ	China	1	542	17	104	421
Beijing Univ	China	2	435	139	221	75
Seoul National Univ	Korea	3	239	56	76	107
Cornell Univ	USA	4	210	108	58	44
Univ of California-Berkeley	USA	5	207	92	59	56
National Taiwan Univ	Taiwan	6	176	64	49	63
Massachusetts Inst of Tech	USA	7	171	44	64	63
Univ of Sci & Tech China	China	8	157	20	87	50
Univ of Illinois at Urbana-Champaign	USA	9	153	70	27	56
Fudan Univ	China	10	140	49	65	26
Nanking Univ	China	11	138	42	68	28
Univ of Mumbai	India	12	136	55	23	58
Pennsylvania State Univ-Main Campus	USA	13	136	70	23	43
Univ of Michigan-Ann Arbor	USA	14	134	52	34	48
Shanghai Jiaotong Univ	China	15	133	8	27	98
Univ of Florida	USA	16	132	71	23	38
Nankai Univ	China	17	128	43	65	20
Univ of Wisconsin-Madison	USA	18	125	74	27	24
The Univ of Texas at Austin	USA	19	122	58	30	34
Univ of California-Davis	USA	20	119	75	29	15
Harvard Univ	USA	21	118	59	48	11
Brigham Young Univ	USA	22	116	52	39	25
Univ of California-Los Angeles	USA	23	116	61	38	17
Zhejiang Univ	China	24	115	9	31	75
China Univ of Sci and Tech	China	25	115	20	68	27
Yonsei Univ	China	26	112	21	28	63
Univ of Virginia-Main Campus	USA	27	106	50	29	27
Texas A&M Univ	USA	28	106	62	18	26
Rutgers Univ	USA	29	105	55	25	25
Univ of California-San Diego	USA	30	103	62	24	17
Virginia Polytechnic Inst and State Univ	USA	31	102	45	23	34
Wuhan Univ	China	32	101	36	34	31
North Carolina State Univ at Raleigh	USA	33	97	31	27	39
Middle East Technical Univ	Turkey	34	95	10	24	61
Stanford Univ	USA	35	94	45	27	22
Indian Inst of Tech (IIT) - Madras	India	36	93	0	14	79
Tianjin Univ	China	37	93	3	24	66
Univ of Minnesota-Twin Cities	USA	38	93	41	23	29
Ohio State Univ-Main Campus	USA	39	92	43	21	28
Indian Inst of Tech (IIT) - Kharagpur	India	40	89	5	29	55
Univ of Arizona	USA	41	88	52	20	16

Source: Data from NORC(2008) based upon the NSF Survey of Earned Doctorates

Table 2 - Using biographic information to verify the quality of the name matching

<b>With Chinese name</b>	
<i>Right Match</i>	90.21% are educated in China
<i>Wrong Match</i>	9.79% not educated in China (5.23% Taiwan, 3.49% US)
<b>Educated in China</b>	
<i>Right Match</i>	91.19% have a Chinese name
<i>Wrong Match</i>	8.81% do not have a Chinese name

Table 3 - Descriptive Statistics

<b>Variable:</b>	Students with a Chinese name (n=4,404)				other students (n=27,656)			
	$\mu$	$\sigma$	Min	Max	$\mu$	$\sigma$	Min	Max
Any publications	0.280	0.449	0	1	0.232	0.422	0	1
# of publications	0.519	1.092	0	13	0.378	0.862	0	17
Weighted # of publications	2.867	6.158	0	72.637	2.148	5.038	0	75.033

Note: first-authored publications only.

Table 4: Main results

	Grad student productivity (I)	Grad student productivity (II)	Grad student productivity (III)
Chinese student	1.417*** [0.0104]	1.092*** [0.149]	1.340*** [0.105]
Late*Chinese Student		0.730*** [0.204]	
Chinese Advisor			2.118*** [0.454]
Chinese Advisor*Chinese Student			2.956*** [0.7211]
Constant	1.68*** [0.150]	1.734*** [0.151]	1.683*** [0.154]
Graduation Year FE	yes	yes	yes
University FE and Subfield FE	yes	yes	yes
R-squared	0.06	0.06	0.06

Notes: 31,162 Observations. OLS estimates. Robust standard errors in brackets.

\* significant at 10%, \*\*significant at 5%, \*\*\* significant at 1%

Table 5 - Results with the biographical subsample

	Grad student productivity (I)	Grad student productivity (II)	Grad student productivity (III)
Chinese name	1.231*** [0.229]		0.425 [0.415]
Undergraduate in China		1.405*** [0.441]	1.029*** [0.415]
Constant	1.438** [0.325]	1.423*** [0.325]	1.042*** [0.325]
Graduation year fixed effects	yes	yes	yes
Subfield fixed effects	yes	yes	yes
University fixed effects	yes	yes	yes
R-squared	0.07	0.05	0.05

Notes: 6792 Observations. OLS Estimates. Robust standard errors in brackets.

\* significant at 10%, \*\*significant at 5%, \*\*\* significant at 1%