University of Maryland, Department of economics
Master of Science in Applied Economics
ECON 644: Empirical Analysis II (Introduction to Economic Models)

TITLE PAGE

Replication: The Value of U.S. College Education in Chinese Labor Markets

Mihir N Madhekar

Department of Economics, University of Maryland

ECON644: Empirical Analysis II: Introduction to Economic Models

Dr Hossein Abbasi

May 17, 2024

Summary

This research aims to assess the replicability of Chen's (2024) study on the impact of a U.S. college education on employment call-back rates in China, particularly in the fields of business and computer science. Utilizing the original dataset from Chen's large-scale field experiment, this study applies both linear probability and Logit models to test the robustness of the original findings. The analysis extends to examining the role of posted wages and the influence of cultural perceptions on employer decisions, further breaking down the posted wage by chosen major to identify significant trends. Initial findings suggest that applicants with U.S. degrees face a statistically significant disadvantage in call-back rates compared to their counterparts with Chinese degrees, with the gap varying by job sector and wage level. This paper contributes to the ongoing discourse on the value of foreign degrees in global labour markets by providing an understanding of the challenges and opportunities faced by U.S.-educated individuals in China.

Introduction

Chen studies the value of a US college education in global labour markets, with a specific focus on China. The US welcomes a million international students for education each year, half of whom are undergraduates. These undergraduate students spend nearly $45 billion on tuition and living expenses within the US and contribute significantly to the intellectual and cultural diversity of American universities. Even though these students can attain an undergraduate education in the US, the majority are required to leave the country due to various reasons, primarily visa limitations, and therefore choose to work overseas. In this study, Chen focuses on the impact of having a US education on global labour markets at the offer stage, with a specific focus on China, given that it accounts for nearly one-third of the international undergraduates in the US.

Chen conducts a large-scale field experiment, where she gauges employer perceptions of college education from the US through a vignette choice experiment method. Along with this, she also uses administrative data on enrolments and actual college resumes to generate fictitious profiles and submits over 27,000 job applications in three major Chinese cities. For this study, Chen focuses on the 'call-back' received at the offer stage, where a call-back is defined as a personalized and positive contact from a potential employer.

Chen identifies the scale of the higher education market for Chinese students and the distinction in opportunities and scale of offerings. There is a growing market for international students within US universities, aided by two major identified factors. Firstly, due to large appropriation cuts by the government for public universities, there is a strong motivation to recruit out-of-state students who pay a higher tuition fee. Bound et al. (2020) estimate that a 10% reduction in state appropriations is associated with a 12% increase in foreign student enrolment at public research universities. Further, due to flexible exchange rates and easier access to credit, US universities have become financially accessible for Chinese families.

Chen designs the experiment in two parts. Firstly, she chooses universities from both the US and China and divides them as per required standardized test scores and internationally accepted rankings. The study is focused on business and computer science majors, given that they account for half of all Chinese undergraduate students who entered the US in the fall of 2014. For each target job, Chen sends four fictitious but realistic job applications, the characteristics of which are randomized. Chen randomly varies the country and selectivity of the university and randomly varies other factors namely, high school, work experience, gender, and job season, and controls for posted salary and firm ownership.

My paper investigates whether the analytical results of Chen (2024) are replicable and whether the experimental design and outcomes hold under other models. For this study, I use the provided dataset from Chen (2024), given that the data was collected as part of a controlled experiment and I want to test its validity against other models.

Chen has focused on multiple groups of control variables to test the impact of US education on call-back rates, but I want to focus mainly on the occupation and prior US work experience. The other groups of control variables are firm and job posting specific, while occupation and work experience are variables that are student-related. The primary research question for this replication will be, ***"the impact of an undergraduate degree from the US on call-backs for employment in China, with a focus on the choice of the student's major and the posted salary."***

I focus on the choice of the student's major primarily because this is under the student's control. The decision to pursue an international education at a higher total cost needs to be taken after due consideration, the biggest being the availability of opportunities and return on investment post-completion. These sets of variables allow us to isolate the true impact of getting an education in the US while controlling for factors that are directly influenced by the student. The results can be further extrapolated to allow the student to have information for more concrete decision-making. For whether or not to take work experience in the US immediately after education can be answered by seeing its impact on getting a call-back, depending upon individual comfort level, an appropriate p-value will allow the student to make an informed decision.

For a further robust check, I compare the author's use of the linear probability model using fixed effects with a Logit model. With the call-backs codified appropriately, I aim to test the validity of the results and the proposed hypotheses under other models. The Logit model is specifically designed to handle binary outcomes. Testing it against the linear probability model can help determine if the model provides a better fit for the data, particularly when the dependent variable is dichotomous.

Data and Descriptive Analysis

Chen focuses on the bachelor's program, particularly in business and computer science, as these programs accounted for half of all Chinese undergraduates in the US who entered in Fall 2024. The data is collected using an experiment in which the author has applied for jobs under aliases, varying the demographic and other control variables. As part of the experiment, the author applies for the jobs as per the following table. The applications are made on a large, nationally recognized job board. Column (4) of Table 1 shows the percentage of active full-time opportunities for the particular qualification as a percentage of total full-time jobs. This shows that the requirement for business and computer science graduates encompasses nearly 50% of all open applications.

Occupation Category	Degree Programs	Sample Job Titles	Share of full-time vacancies
Business Accounting, Finance, Banking, Sales, Customer Service, Marketing	Accounting, finance, economics, marketing, business administration	Accountant, billing/payroll spec, business associate, financial analyst, project assistant, sales associate, account manager, marketing spec	0.315
Computer Science Software, Network, IT	Computer science	Web developer, software engineer, testing engineer	0.191

Table 1

Randomisation Structure

For each job, Chen sent four fictitious but realistic resumes. Here, the author varied resume characteristics, including demographics, country, selectivity of institution, etc. Chen has followed the general practice in the audit study literature, based on which two of the four resumes are assigned to universities from China and two are assigned to universities from the United States. The author has randomized the selectivity of the institution based on the selectivity parameters as per the introduction.

Chen has also randomized other resume characteristics, namely, high school attended, work experience in the United States, current work status, and gender. This allows us to control for individual variables and identify correlations. Additionally, the applications are sent across two hiring seasons, spring and fall recruitment. During the spring recruitment cycle, the author also randomly signalled a high score on China's college entrance exam among applicants with U.S. degrees. Chen sent four resumes for each job over two to three days with at least a four-hour gap to reduce the likelihood of suspected linkages. Employers' responses are meticulously tracked via phone, text, or email. And a 'call-back' is defined as personalized positive contact from employers.

Resume Construction

Chen has chosen the U.S. institutions to represent the majority of the Chinese students that are studying in the country. For this purpose, the author has excluded the institutions that enrol fewer than 30 Chinese students in a batch or are unranked as per U.S. News Rankings or rank below 250 in the same. With these criteria, the experimental sample contains 111 institutions from the U.S. that enrolled 72% of all Chinese students that started in fall 2014. Chinese institutions are chosen from the official list of all postsecondary institutions published by China's Ministry of Education. The author excluded specialized and strategic schools and used local institutions that grant bachelor's degrees in Beijing and Shanghai to avoid employers' concern about potential selection based on who leaves their hometown for college and who returns. After comparing the institution ranking and average test scores, Chen is able to positively conclude that her determination of the selectivity criteria is robust. Average test scores tend to increase as the schools go higher in ranking. And the universities that are the most inclusive tend to have a wider distribution of test scores and enrol a higher percentage of students from the lower quartile.

Descriptive Statistics

For the purpose of this paper, the primary variable of interest is whether or not an application receives a 'call-back'. A call-back is defined as a "personalized and positive contact from a potential employer." Chen has identified multiple control variables that have been used in resume construction that allows for varying demographics to identify their isolated impact on receiving call-backs.

Table 2 allows us to get an initial understanding of the data and a broad breakdown of the resume characteristics. We see that the average rate of overall call-back, across a comprehensive sample of over 26,000 applications, is 15%. This is further broken down by the various explanatory variables that we will focus on as part of this study. The variation in resume characteristics has been purposely maintained to be proportional to reduce inter characteristic variation. The primary country of education of the resumes is equally divided between the US and China in line with the accepted standard practices for such an experiment. The proportion of applicants with a 'Business' major is higher owing to the higher number of open jobs for the major, which is supplemented by the information from Table 1. Given that the subject encompasses various sub-majors that are individually considered professions as well, a higher weightage is justified. Figure 1 is a graphical depiction of the average rate of call-backs received as part of the entire experiment.

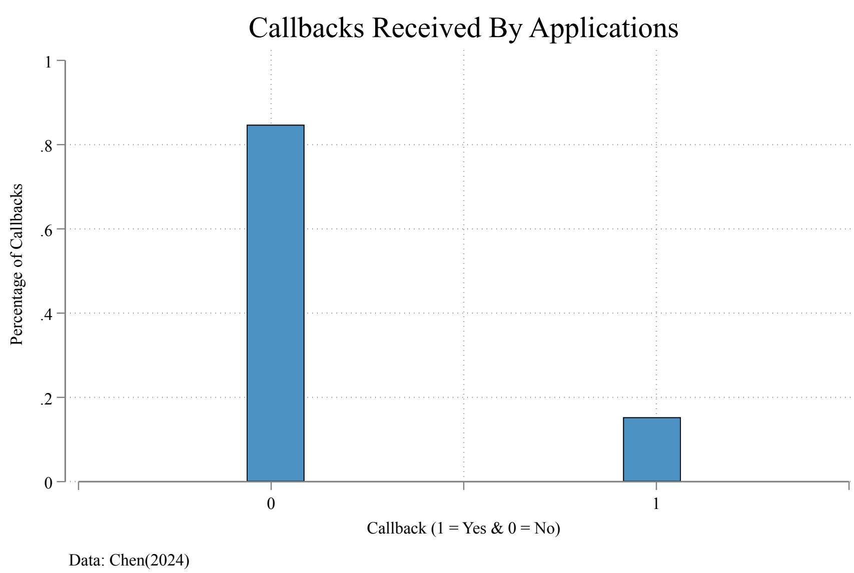

Figure 1: Total call-backs received by applicants.

Note: This includes all the 26,036 applications made as a part of the experiment.

Variables	Resumes Sent	Mean Call Back Rates
Total	26,036	0.150
By Country of Education		
United States	13,018	0.135
China	13,018	0.165
By Selectivity of Institution		
Very Selective	8,689	0.157
Selective	8,636	0.154
Inclusive	8,711	0.140
By Major		
Business	19,108	0.176
Computer Science	6,928	0.078
By Gender		
Female	12,628	0.166
Male	13,408	0.136
By Application Cycle		
Fall	11,340	0.134
Spring	14,696	0.163

Table 2: Summary statistics for 'Call-back'

Note: I have reported only the mean value for the variable of interest and the conditional statistics based on the explanatory variables given that the variable of interest is a binary indicator and the detailed summary statistics do not provide additional information. The same have been reported in the log file.

Table 3 gives us an insight into the explanatory variables considered for this replication. 'Job ID' is used as a key for individual jobs for which 4 separate resumes have been sent as part of the experimental design. 'Business', 'Experience', 'Female', 'Very Selective' and 'Selective' are binary variables that indicate various resume characteristics that have been altered. Chen runs a model as part of the original paper to test the impact of posted salary on the call-backs received. There is a key assumption with regards to the impact of salary. It is perceived that for the jobs with a lower posted salary, the candidate with a US/foreign education would be overqualified and would potentially seek employment elsewhere. The paper tests this assumption by dividing the salary by quartile and testing it with an interaction term.

Variables	Number	Mean	Std Dev	Min	Max
US	26,036	0.500	0.500	0	1
Job ID	26,036	3,335	1,978	1	6,780
Business	26,036	0.734	0.442	0	1
Experience	26,036	0.526	0.499	0	1
Female	26,036	0.485	0.500	0	1
Wage	26,036	15,531	7,306	5,548	55,479
Interview	26,036	0.138	0.345	0	1
Very Selective	26,036	0.334	0.472	0	1
Selective	26,036	0.332	0.471	0	1

Table 3: Summary statistics for explanatory variables

Figure 2 allows us to see the distribution of the resume characteristics across the experiment sample. Given that most of the explanatory variables are binary by design, the graphs allow us to see their distribution as a percentage of the whole sample.

The two key explanatory variables that allow us to manipulate and test the biggest perceived differences include the wage posted and the selectivity of the institution. Figure 3 shows us that for the 6,780 jobs that were applied for, the posted wages have a bi-modal distribution. With the lower wages constituting the majority of the jobs. Given that the jobs were identified for early career professionals, and recent graduates, the expected posted wage for the jobs would be concentrated around a lower number. Further, we identify the possible correlation between the posted wage and the major. The posted wages are higher for ‘Computer Science’ graduates than for ‘Business’ graduates. This wage gap coupled with the higher gross number of job openings for Business graduates can be an indicator of excess supply of students with the major.

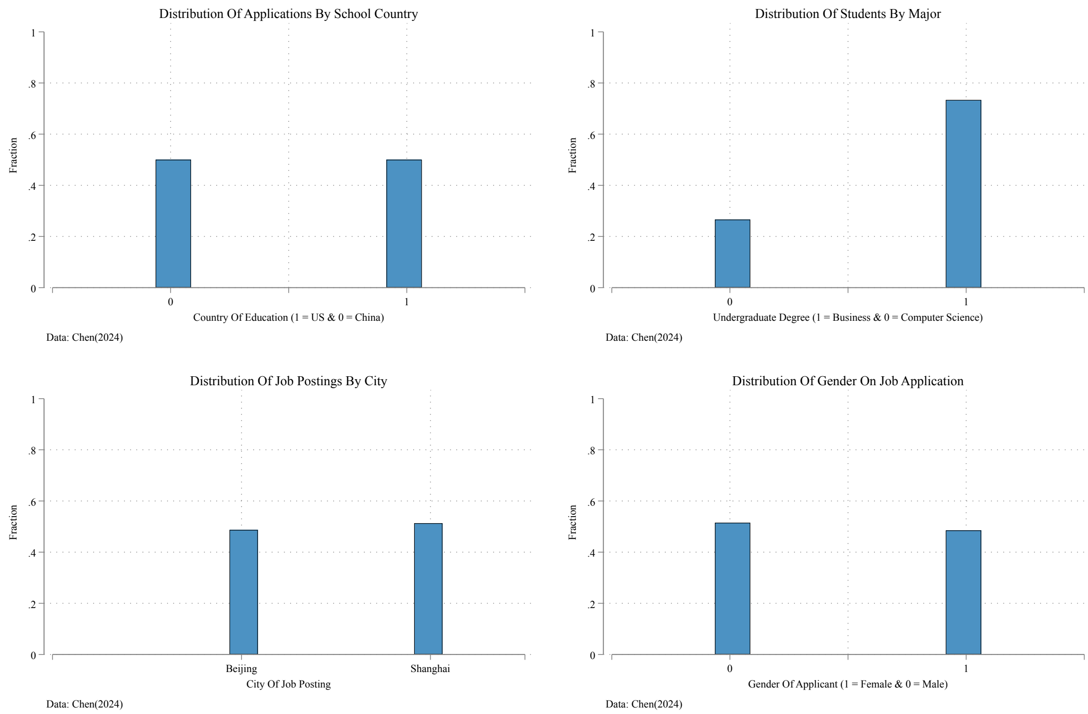

Figure 2: Descriptive statistics for explanatory variables

Note: (From left to right) Panel 1: Distribution by country of education, Panel 2: Distribution of Major, Panel 3: Distribution by City of posting and Panel 4: Distribution by Gender.

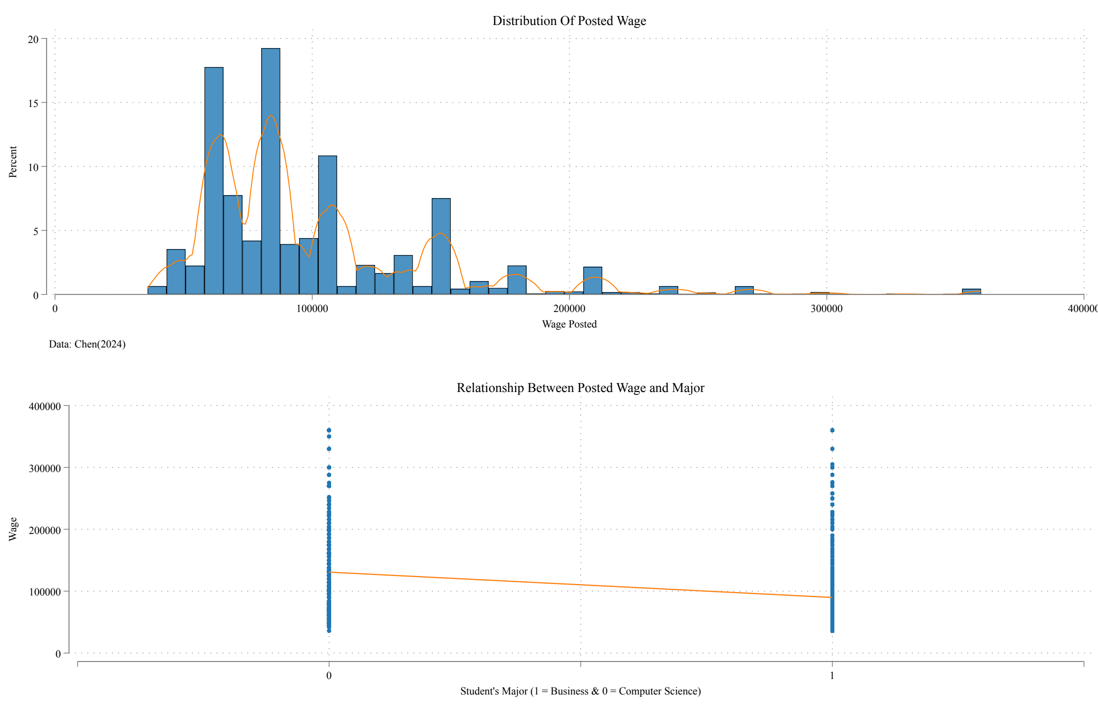

Figure 3: Descriptive statistics for posted Wage and Relationship between Wage and Major

The final key explanatory variable that is to be considered is the classification of the institutions. Chen has classified the institutions to be ‘Very Selective’, ‘Selective’, and ‘Inclusive’ based on different criteria for each country. For the US, the institutions are classified based on their ranking in the US News Rankings, Top 50 being Very Selective, Top 250 being Selective, and the rest being inclusive. And for China, the selectivity of the institution is dependent on its inclusion in either the 985 project or the 211 project.

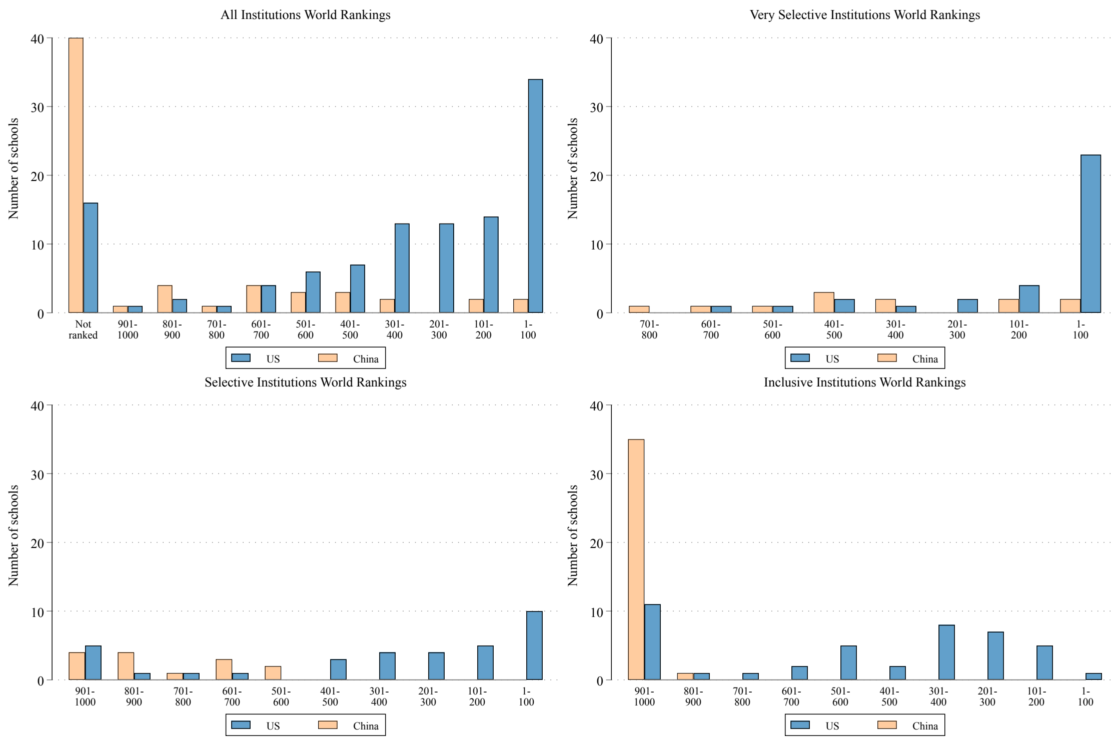

Figure 4: Selectivity and Rankings of Institutions

Note: (From left to right) Panel 1: All Institutions, Panel 2: Very Selective Institutions, Panel 3: Selective Institutions and Panel 4: Inclusive Institutions

Figure 4 allows us to verify the classification given that the institutes from China may not necessarily conform to global rankings or may not be included in the same. We see that the institutes that are a part of the Project 985, are concentrated among the Top 200 and Top 300 – 400 on world rankings. Whereas, the institutes that are part of Project 211 are ranked among Top 500 and below. This is roughly representative of the selection criteria of the author. Last, although sample U.S. institutions are more highly ranked than Chinese institutions on average, the distribution of world rankings is similar to that of all ranked schools. For the choice of the institution sample, the author has used the actual enrolment of Chinese students in 2014 as sample weights, resulting in only the relevant institutions being included in the sample.

This descriptive analysis allows us to get an understanding of the dataset and the relevant explanatory variables. The distribution of the wages and classification of the universities allows us to verify the underlying assumptions prior to running the model.

Empirical Analysis

Chen estimates the differential call-back rate between the United States and China across the paper using variations of a linear probability model. The model is defined as follows:

$$Callback_{ij} = \beta(U.S.degree_{ij} + Other\ resume\ charactersitics_{ij}\theta + \omega_j + \epsilon_{ij}$$

Where the dependent variable is an indicator for applicant i receiving a call-back for job vacancy j, and the independent variable of interest, $\mathrm{U.S\ Degree}$ is an indicator for whether a U.S. Institution is listed on the job application. Coefficient β is the marginal effect of having a U.S. College education, relative to a college education in China on the probability of having a call-back; ω_j represents vacancy fixed effects that are present in the preferred specification. Standard errors are clustered at the vacancy level at all times.

We can verify the results of the linear probability model using the Logit model with marginal effects to test the validity of the results and compare coefficients. Table 4 is the primary regression that shows the main results of the United States-China gap in call-back rates estimated using the linear probability regression model established above and with a Logit model.

Table 4 shows the main results that will set a benchmark, allowing us to further analyse the call-back gap using other resume characteristics. Column 1 in Table 4 has no other dependent variables and Column 2 adds controls for work status, gender, application cycle, labour market, and self-statement template. The initial regression shows that there is a difference of 3 percentage points between the call-back rates of Chinese and U.S. education, which is statistically significant at every conventional level. The difference is consistent with the added additional control variables and is robust after being verified using a Logit model.

To accurately interpret the true effect of the regression, we use a baseline call-back rate. We have considered education from China to be the omitted variable and we consider the average call-back rate for this variable as the baseline. Across the current models, the baseline call-back rate is 0.165, i.e., 16.5%, for job applicants with a degree from China. From the regression results, we can see that the average call-back rate for an applicant with a degree from the United States is 3 percentage points lower, i.e., 13.5%. This tells us that on average, an applicant with a U.S. Degree is 18.2% less likely to receive a call-back compared to an applicant with a degree from China.

	All	All	Logit	M. Effects	Logit	M. Effects
	(1)	(2)	(3)	(4)	(5)	(6)
U.S. Degree	- 0.030*** (0.003)	- 0.030*** (0.003)	- 0.238*** (0.026)	- 0.030*** (0.003)	- 0.245*** (0.027)	- 0.030*** (0.003)
Working Full Time		- 0.023** (0.010)			- 0.332* (0.130)	- 0.041* (0.010)
Female		0.030*** (0.007)			0.241*** (0.056)	0.030*** (0.007)
Spring Job Season		0.029*** (0.007)			0.233*** (0.058)	0.029*** (0.007)
China Call Back Rate	0.165	0.165	0.165	0.165	0.165	0.165
Observations	26036	26036	26036	26036	26036	26036
Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001						

Table 4: Call-back Regressions (Controlling: Full Time Experience, Gender, Application Season)

Note: Dependent Variable is the indicator for receiving call-backs. Degree from China is the omitted variable. Column 2 controls for the above mentioned variables. Standard errors are clustered at the vacancy level.

This is statistically significant at the 1% level even after adding additional control variables. From Columns 2 and 6 of Table 4, we can further identify that gender and job application cycle are statistically significant at the 1% level and have a positive impact on the rate of call-backs received. At a cursory level, a female applicant with a Chinese education, without prior work experience, who applies in the Spring cycle has the highest potential call-back rate.

Having a baseline of the results, we can include name and vacancy fixed effects, which absorb controls that are randomized across jobs. Table 5, Column 1 includes the baseline regression with name and self-evaluation fixed effects. We further break down the regression results by the choice of major. This is one of the key variables that we are focusing on for the purpose of this replication. Column 2 of Table 5 follows the specification from Column 1 but shows the results for 'Business' majors and similarly Column 3 shows the results for 'Computer Science' majors.

	All (1)	Business (2)	Comp. Sci (3)
U.S. Degree	- 0.030*** (0.003)	- 0.032*** (0.004)	- 0.026*** (0.006)
China Call Back Rate	0.165	0.192	0.091
Observations	26036	19108	6928
Vacancy Fixed Effects	Yes	Yes	Yes
Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001			

Table 5: Call-back Regressions With Fixed Effects and Majors

Note: Dependent Variable is the indicator for receiving call-backs. Degree from China is the omitted variable. Column 1 to 3 include fixed effects for names and self-statement templates.

	All (1)	All (mfx) (2)	Business (3)	Business (mfx) (4)	Comp Sci. (5)	Comp Sci. (mfx) (6)
U.S. Degree	- 0.242*** (0.027)	- 0.030*** (0.003)	- 0.221*** (0.028)	- 0.032*** (0.004)	- 0.365*** (0.079)	- 0.026*** (0.006)
China Call Back Rate	0.165	0.165	0.192	0.192	0.091	0.091
Observations	26036	26036	19108	19108	6928	6928
Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001						

Table 6: Call-back Regressions Logit Model With Marginal Effects

Note: Dependent Variable is the indicator for receiving call-backs. Degree from China is the omitted variable.

From the above results of Table 5, we see that applicants with a U.S. education receive a call-back 3.2 percentage points less for Business jobs and 2.6 percentage points less for Computer Science jobs when compared with a Chinese education. This shows that applicants with a U.S. education are 16.6% less likely to receive a call-back for a business job and 28.5% less likely to receive a call-back for a computer science job when compared with an applicant with a Chinese education. These results are statistically significant at a 1% level and economically significant given that the average disparity is 22.5%.

I have also used a Logit model with marginal effects, presented in Table 6, to verify the outcomes of the linear probability model. The results and the coefficients are equal across the entire model.

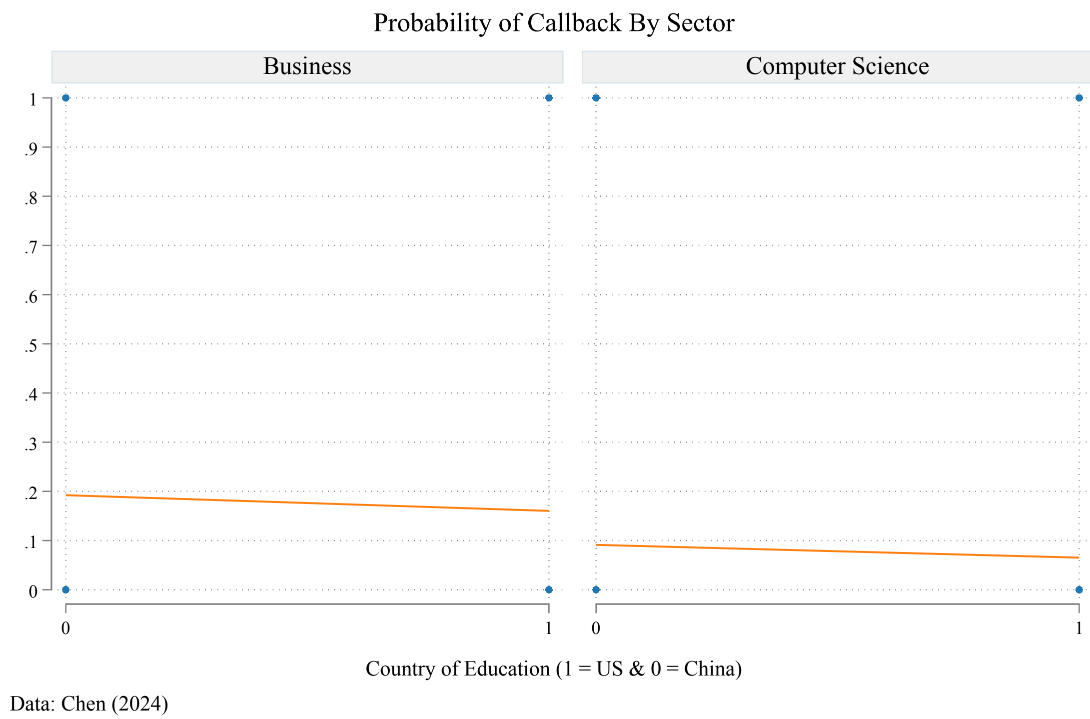

Figure 5: Call-back rates by Job Sector

The regression results can be visualized as per Figure 5. We see a decline in call-back rates from 0 indicating education in China to 1 indicating education in the US. This is true for both the job sectors seen individually and allows us to visualize the intensity of the effect.

The results from the above regressions are indicative of a United States-China call-back gap that is evident and statistically significant. This is indicative of a difference in employer perception about applicant quality based on cues from educational background. Chen through this experiment has strategically differentiated the resumes, but there are a limited number of personal and professional characteristics that can be revealed through a resume at first glance. But, with the significant gap in call-back rates, it can be assumed that there are inherent employer perceptions associated with education from the United States. Although the author explores four distinct mechanisms with empirical evidence to explain the patterns, we will focus on 'posted wage' as a part of this study.

Impact Of Posted Wages

The cost of employment consists of multiple associated costs for the employer, which include the cost of interviewing candidates, scheduling and conducting preliminary screening, and the cost of training. The employers also have to consider the retention costs for the employee. If they are more likely to find employment elsewhere, the employer does not have any incentive to spend on their recruitment.

There is a cultural perception among Chinese employers that candidates with foreign education would prefer to and find better opportunities elsewhere. This is further exacerbated by the perception that the candidates with education from local universities would not have the same opportunities. The US-China call-back gap can have a possible explanation in the premise that the candidates with education from the US would be in a position to secure better opportunities out of the country and hence would prefer to do so. Given that they have access to the US labour market, at least for their initial period, the candidates would prefer to take up opportunities from the same.

Table 7 shows us the impact of the posted salary on getting a call-back. The omitted variable is the education from China. Hence, the regression tells us about the probability of getting a call-back given that the candidate has education from the US and depending on the posted salary for the job. Chen has used the posted salary as a proxy for the quality of the job. Even though this may be a fair approximation, there is a difference between the distribution of the posted salaries by major as well. This has not been expressly focused upon by the author and has been explored further as a scope of this paper.

When we see the trend of the call-back gap between the US and China, we see that the gap reduces as there is an increase in the posted salary. Table 7 presents the combined effects of the posted salary on call-back rates for both majors. Column 1 is a baseline regression for the impact of salary and Columns 2 to 10 break down the impact by deciles. The median salary is at 84,000 RMB or about $12,000 USD. The regressions account for name and self-evaluation fixed effects as a standard. The coefficients for the variables US degree, show a decreasing trend, that is indicative of the reducing US-China call-back gap as the salary increases. For the baseline, at the mean posted salary, a candidate with a US degree has a 6.5% lesser probability of receiving a call-back when compared with a Chinese candidate.

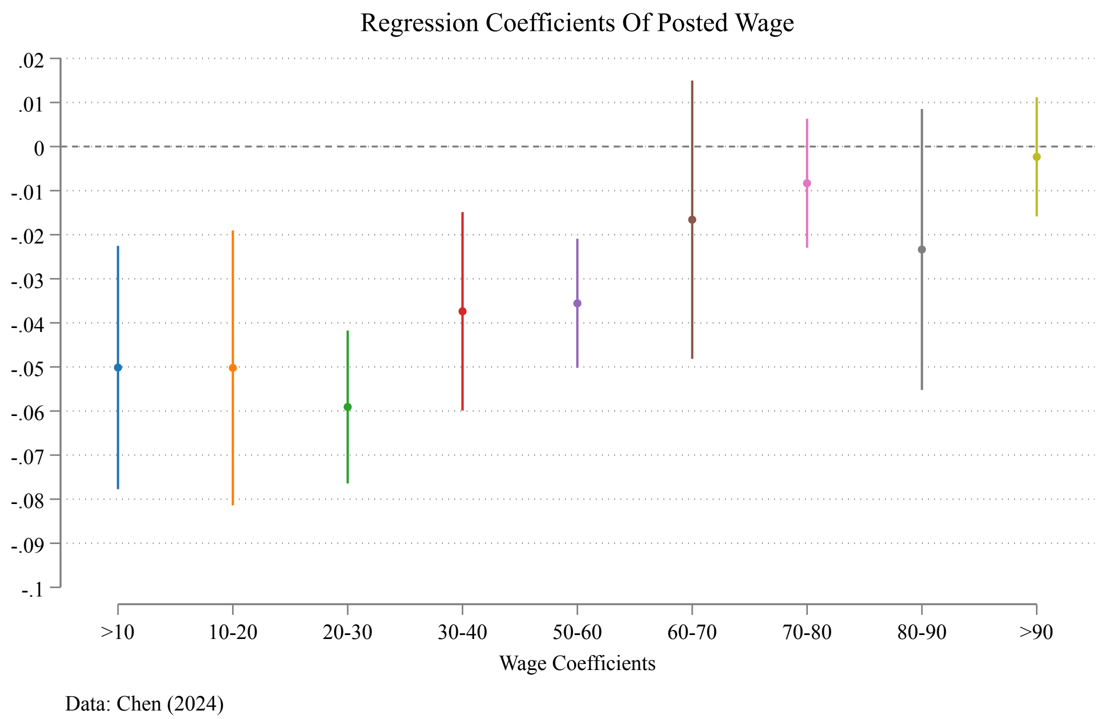

Figure 6: Call-back regression coefficients by posted wage deciles

Note: Each dot represents the coefficients of the regression from Tables 7. The lines represent the 95% confidence interval of the coefficient. The regression follows the same specifications as Column 1 of table 5. It has fixed effects for names, self-evaluation and vacancy. The standard errors are clustered at the vacancy level. The 40 – 50th decile is omitted as it has no values.

This is also seen when we plot the coefficients. We see that the gap is reducing but is never positive. It is interesting to note that the number of jobs is concentrated in the lower deciles, due to them being entry-level jobs, and the call-back gap is the most significant in the 20–30 decile. The employers can be presumed to be hiring students in greater numbers and would prefer locally educated students who may not have extensive financial burdens of education loans and would be willing to work at the lower salary.

At the higher deciles, and by assumption, quality of job, the mean call-back rate for students with an education from China is 9.5%, which is in itself an indicator of the exclusivity or selectivity of the job. Given that this, the gap of 0.2 percentage points, would be justified, with the assumption that at the higher salary option, the graduate would be indifferent between working in China or elsewhere and similarly the employer is willing to spend money to attract talent.

	Baseline (1)	>10 (2)	10-20 (3)	20-30 (4)	30-40 (5)	50-60 (6)	60-70 (7)	70-80 (8)	80-90 (9)	>90 (10)
U.S. Degree	- 0.065***	- 0.050***	- 0.050**	- 0.059***	- 0.037**	- 0.036***	- 0.017	- 0.008	- 0.023	- 0.002
	(0.007)	(0.014)	(0.016)	(0.009)	(0.011)	(0.007)	(0.016)	(0.007)	(0.016)	(0.007)
China Call Back Rate	0.165	0.224	0.205	0.219	0.207	0.178	0.141	0.123	0.121	0.095
Observations	26036	1680	1052	4460	2328	5932	1148	4148	824	4464

Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001

Table 7: Call-back regression by Salary deciles

Note: The regression follows the same specifications as Column 1 of table 5. It has fixed effects for names, self-evaluation and vacancy. The standard errors are clustered at the vacancy level. The 40 – 50th decile is omitted as it has no values.

	Baseline (1)	>10 (2)	10-20 (3)	20-30 (4)	30-40 (5)	50-60 (6)	60-70 (7)	70-80 (8)	80-90 (9)	>90 (10)
U.S. Degree	- 0.065***	- 0.051	- 0.117	- 0.126***	- 0.093**	- 0.054*	0.004	- 0.016	- 0.019	0.001
	(0.007)	(0.036)	(0.072)	(0.030)	(0.031)	(0.019)	(0.024)	(0.010)	(0.019)	(0.008)
China Call Back Rate	0.165	0.093	0.152	0.212	0.125	0.160	0.078	0.063	0.080	0.059
Observations	26036	172	92	424	304	1000	360	1516	476	2584

Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001

Table 8: Call-back regression by Salary deciles – Computer Science Majors

Note: The regression follows the same specifications as Column 1 of table 5. It has fixed effects for names, self-evaluation and vacancy. The standard errors are clustered at the vacancy level. The 40 – 50th decile is omitted as it has no values.

	Baseline (1)	>10 (2)	10-20 (3)	20-30 (4)	30-40 (5)	50-60 (6)	60-70 (7)	70-80 (8)	80-90 (9)	>90 (10)
U.S. Degree	- 0.065***	- 0.051***	- 0.047**	- 0.053***	- 0.029*	- 0.032***	- 0.025	- 0.004	- 0.035	- 0.007
	(0.007)	(0.015)	(0.017)	(0.009)	(0.012)	(0.008)	(0.007)	(0.010)	(0.029)	(0.012)
China Call Back Rate	0.165	0.239	0.210	0.220	0.219	0.182	0.170	0.158	0.178	0.145
Observations	26036	1508	960	4036	2024	4932	788	2632	348	1880
Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001										

Table 9: Call-back regression by Salary deciles – Business Majors

Note: The regression follows the same specifications as Column 1 of table 5. It has fixed effects for names, self-evaluation and vacancy. The standard errors are clustered at the vacancy level. The 40 – 50th decile is omitted as it has no values.

Chen, in the original paper, has included an interaction term to test to the impact of the US education and its interaction with the posted wage. As a part of this replication, I have added interaction terms across the entire regression of Table 7, the results of which are in Appendix Table 1. The interaction terms for the 20 – 30 decile and the above 90 decile are statistically significant. Although the other interaction terms provide additional information, due to the nature of the data, and the setup of the original regressions, we are able to capture the effect of the interaction terms within the original regression. The coefficients for each of the deciles from Table 7 align with the interaction coefficient values for ß1 + ß3.

When we break down the impact of the posted wage by the sector, we are able to see distinct trends for both the Computer Science and Business fields. It should be noted, that even though not all the coefficients are statistically significant when we break down the posted salaries by sector, it allows us to get an idea of the trend, and for this purpose, I have extended the conventional limits of statistical significance to include all the salary deciles.

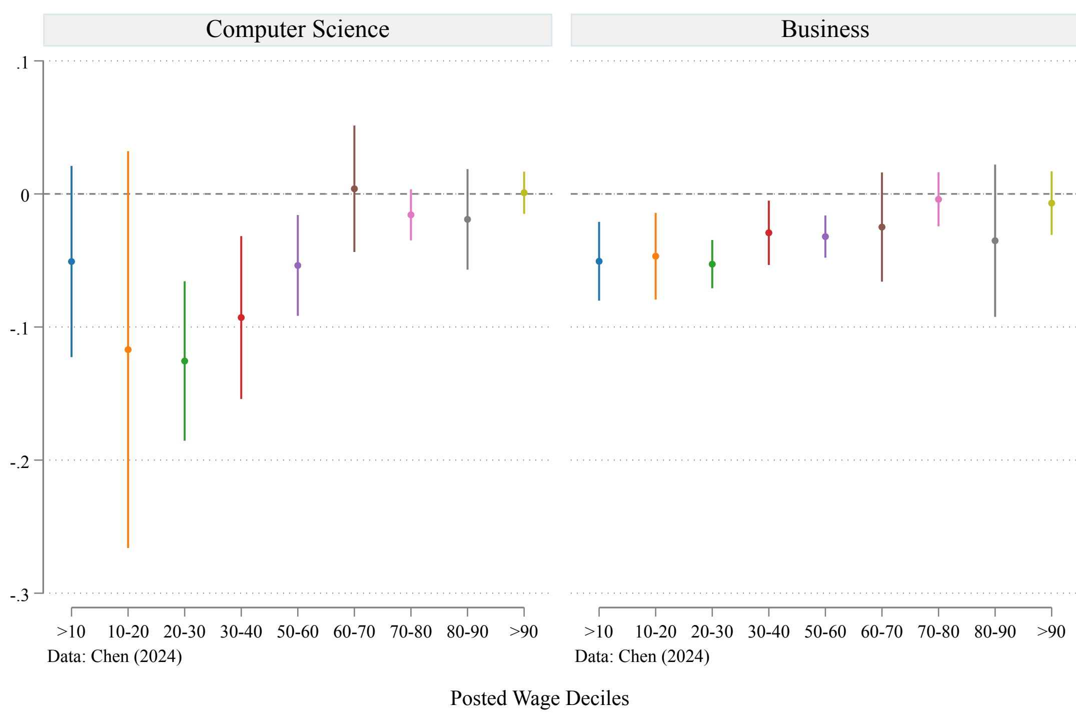

Figure 7: Call-back regression coefficients by posted wage deciles – by sector

Note: Each dot represents the coefficients of the regression from Tables 8 & 9. The lines represent the 95% confidence interval of the coefficient. The regression follows the same specifications as Column 1 of table 5. It has fixed effects for names, self-evaluation and vacancy. The standard errors are clustered at the vacancy level. The 40 – 50th decile is omitted as it has no values.

For the job postings that are from the Computer Science sector, we see the trend following the overall trend across postings. The gap is maximum at the 20–30 decile at 12.6 percentage points and is highly statistically significant, and interestingly shows a positive gap at the 60–70 decile with 0.4 percentage points. The range across the deciles is 13, which is significantly higher than the range for the combined distribution at 4.8 percentage points.

This trend further tells us that even though for the entry-level jobs at the lower deciles, graduates with a local education are preferred, as the posted salaries increase, the preference is shifting towards graduates from the US, with there being positive coefficients at the 60–70 decile and above the 90th decile. We can attribute this to the potentially specialized nature of the education that may be offered in US universities in the fields of Computer Science. Furthermore, for the higher deciles, mean rate of call-backs for graduates with a degree from China are in the range of 7.8% to 5.9%, which shows the selective nature of hiring, where having a foreign education may serve as a significant distinguishing factor.

When we see the trend for the Business sector jobs, the US-China call-back gap follows the similar overall pattern, but is not as extreme. The range is 4.6 percentage points which is fairly consistent with the overall trend. For the Business sector postings, we see no positive gaps although at the highest deciles, the gap is 0.7 percentage points, which is economically low. The most interesting aspect of the Business sector is the call-back rate for graduates with an education from China. Even at the highest decile, the mean call-back rate is 14.5%, which is closer to the overall mean rate and is significantly higher than the 5.9% at the highest decile of the Computer Science postings. This can be attributed to the higher number of job postings, and the increased demand for the major.

After having seen the empirical results for the impact of foreign education while focusing on the choice of major and the posted salaries, we can see that the trend of difference in call-backs is evident throughout the choice of major and the posted salary.

Below the median salary, even though the average call-back is fairly high, the call-back gap is high as well. This would make it exceedingly difficult for entry-level graduates who will seek employment after having done their undergraduate studies in the US. Further, the call-back gap is a pressing concern for people with Computer Science as their major. With the average call-back rate for the US Computer Science graduate being 6.5% compared to a graduate from China, and the gap being over 10 percentage points at the lower deciles, the job market in China may not be highly suitable. On the other hand, the call-back rate for entry-level positions at the lower salary deciles for Business majors is higher than the average, even with the call-back gap accounted for.

Chen further builds on these results using three more mechanisms, studying the Lack of Information on U.S. Colleges, Selection on Unobserved Quality, and the Local Employer Network. These mechanisms have not been included as part of this paper.

Chen further builds on these results using three more mechanisms, studying the Lack of Information on U.S. Colleges, Selection on Unobserved Quality and the Local Employer Network. These mechanism have not been included as a part of this paper.

Concluding Remarks

The allure of a foreign higher education has been apparent in the Eastern world since air travel became cost-effective. This has attracted students from across the continent to invest time, energy, and money to attempt to reap the benefits of these opportunities. In the past, getting an education from the United States was considered the ultimate intellectual achievement for a student, and the benefits were apparent. However, current research shows that not enough is known about the true value of an education in the US in the labour market of other countries. Chen conducts a large-scale field experiment to identify the impact of having a US education on getting entry-level opportunities within the Chinese job market. Her original study is accompanied by a survey of employers within China to gauge their perception of candidates with education from the US. As part of this replication study, I focus on two central mechanisms that attempt to explain the disparity in opportunities for students with an education from the US and China.

The results from the original study tell us that there is a preference for local talent within China, holding all other demographic differences constant. We find that the US-China callback gap is consistent across the two broad sectors and persists at all salary levels. Chen originally broke down the posted wages by quartile, but I have used deciles to get a more granular idea of the trend and further broken it down by sector. This has allowed me to observe varying trends across both sectors that are not apparent from the combined results. At the higher deciles, we see that the gap is almost negligible for technical roles, and we see a positive gap for the 60th–70th decile for Computer Science majors as well. Although it is true that having a US education is not an immediate benefit in the Chinese labour market, at higher salaries, at least, it is not an impediment. Chen has further used other mechanisms to explain the callback gap, such as the ownership of the firm and language requirements.

I have focused on the mechanism that, in my opinion, hold more importance in the decision, but the original study is more comprehensive and explains the impact of the US Education from multiple facets. Although the study does utilise multiple parameters, there is not a lot of cross sectional analysis that has been done within the paper. Given the nature of the experiment and the controls used, the study could be further expanded to identify precise manipulations to a resume that would give the highest callback for a particular job. The dataset has a potential for developing an algorithm to further identify the 'Most Ideal Candidate / Resume' based on the overall industry acceptability.

One of the most important policy implications of this paper is that US universities with a majority of their students being international students should make an active effort to assist these students in transitioning from American education to the Chinese labour market. Given the current trends with immigration and job opportunities, students tend to prefer to return to their primary domicile to work, and assistance from university career services internationally will benefit the students and incentivize more students to attend university in the US.

References

Bertrand M, Mullainathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labour market discrimination. Amer. Econom. Rev. 94(4):991–1013.

Bound J, Breno B, Khanna G, Turner S (2020) A passage to America: University funding and international students. Amer. Econom. J. Econom. Policy 12(1):97–126

CCG and Zhaopin (2017) Report on employment and entrepreneur- ship of Chinese returnees. Technical report, Center for China and Globalization, Beijing.

Chen M, Howell J, Smith J (2020) Best and brightest? International students' decision to attend college in the U.S.. Working paper. Industrial Relations Section Working Paper No. 640, Industrial Relations Section, Princeton, NJ.

Chen M, Liu E, Shang Y (2023) Income and geographic segregation across colleges in China. Mimeo, New York.

Farber HS, Silverman D, von Wachter T (2016) Determinants of call- backs to job applications: An audit study. Amer. Econom. Rev. 106(5):314–318.

Farber HS, Silverman D, von Wachter T (2017) Factors determining callbacks to job applications by the unemployed: An audit study. Russell Sage Foundation J. Soc. Sci. 3(3):168–201.

Gaddis SM (2017) An introduction to audit studies in the social sciences. Audit Studies: Behind the Scenes with Theory, Method, and Nuance (Springer, Berlin), 3–44.

Hastings J, Neilson CA, Zimmerman SD (2015) The effects of earnings disclosure on college enrolment decisions. NBER Working Paper No. 21300, National Bureau of Economic Research, Cam- bridge, MA.

Kline P, Walters C (2019) Audits as evidence: Experiments, ensembles, and enforcement. Preprint, submitted July 18, https:// arxiv.org/abs/1907.06622

Notman, N., & Woolston, C. (2020). Fifteen to one: How many applications it can take to land single academic job offer. Nature, 584(7820), 315–315. https://doi.org/10.1038/d41586-020-02224-5

U.S. News (2017) Global University ranking and National University ranking. https://www.usnews.com/education.

Data Appendix

The original dataset consists of three separate datafiles.

1. full_final – This is the primary data for the study and consists of the randomisation of the resumes and the received responses.
2. Survey_Data – This dataset consists of the survey responses received as a part of the supplementary survey conducted by the author during this study.
3. Ratings – This is a dataset that consists of the ratings of the universities chosen for this study to determine their inclusivity and selectivity

For the purpose of this replication we will be using a modified version of the 'full_final' and the original 'Ratings' data.

Final Data – Modifications

The entire dataset for the resume randomization and the responses received is categorical data that has either been codified within the data file or codified previously. It contains the following variables:

- job_id: unique job identifier
- city: city of the job
- sector: sector of the job
- business: whether the job is for business degrees
- lmarket: labor market of the job
- us: whether the school listed on job application is an US school
- exp: experience type of the job applicant
- *hs: whether signalled elite high school on job application*
- female: whether the job applicant is a female
- name: unique name identifier
- se: unique identifier for self evaluations (which is required on job application template)
- ed_country: the country of the school
- req_en: English requirement of the job
- wage_mid: posted wage midpoint
- wage_mid_d: posted wage midpoint in USD

- wave: wave id of the experiment
- callback: whether the application received a callback
- interview: whether the application received a callback to request for an interview
- vs: whether the school is in the very selective group
- selective: whether the school is in the selective group
- inclusive: whether the school is in the inclusive group
- *cdgrad: the post-graduation career type of the job applicant*
- *wh_type2: unique id for work history type*
- *ctype_cn: whether the firm is a Chinese owned firm*
- *ctype: whether the firm is foreign owned*
- *range: group identifier for school's average test score percentile among admitted students*
- *test: whether signaled entrance exam school on job application*

Note: The variables that are in grey have been dropped from the Final Data as they are not used for the analysis.

After having the selected variables, the following modifications have been performed on the dataset:

1. Encoding Name, SE (self-evaluation), City and Labour Market variables and generating new encoded variables
2. Dropping observations where the city is Guangzhou, as we are focusing only on Beijing and Shanghai
3. Creating a binary variable for the second wave of applications 'wave2'
4. Adding posted wage deciles for mechanism comparison pay10, pay1020, etc
5. Creating a function to make interaction terms for any variable with:
    a. Degree from the US
    b. Very Selective University
    c. Selective University
6. Adding labels to the variables for easy identification

The dataset 'Ratings.dta' has been taken as it is as no further modifications are required to the dataset.

Appendix

	Baseline	>10	10-20	20-30	30-40	50-60	60-70	70-80	80-90	>90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
U.S. Degree	- 0.065***	- 0.029***	- 0.029**	- 0.024***	- 0.030*	- 0.028***	- 0.031***	- 0.034***	- 0.031***	- 0.036***
	(0.007)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
US X >10		- 0.021								
		(0.014)								
US X 10-20			- 0.022							
			(0.016)							
US X 20-30				- 0.035***						
				(0.010)						
US X 30-40					- 0.008					
					(0.012)					
US X 50-60						- 0.008				
						(0.008)				
US X 60-70							- 0.015			
							(0.017)			
US X 70-80								- 0.026**		
								(0.008)		
US X 80-90									- 0.007	
									(0.016)	
US X >90										0.034***
										(0.008)
China Call Back Rate	0.165	0.224	0.205	0.219	0.207	0.178	0.006	0.006	0.005	

Standard errors in parentheses | * p<0.05, ** p<0.01, *** p<0.001

Appendix Table 1: Call-back regression by Salary deciles including Interaction terms