Study Abroad during College: Comparison between China and the United States

Jing Li
Teachers College, Columbia University

Fei Guo
Institute of Education, Tsinghua University

You You
Graduate School of Education, Peking University

Study abroad can be a life-changing experience, but evidence of its effectiveness is mixed. We examine the experience of studying abroad at colleges in the US and China, which are the largest receiving and sending countries of international students respectively. Using data from two comparable national surveys that follow the same design, we estimate causal effects by matching students who studied abroad during college to those who did not, based on their propensity to study abroad. First, we survey student profiles to better understand who studied abroad. We found that parental education and urbanity makes difference in participation. Second, we examine the impact of study abroad on student academic achievement, using multivariate regression and propensity score matching. Finally, we exploit the matched samples to examine impact heterogeneity by student background. We find a positive, statistically significant but small impact on student academic achievement in both countries, with a higher impact for American students.

Introduction

The United States, a traditional receiving country of international students, has started to send more students abroad. One out of ten students now have overseas learning experience during college. In some colleges and universities, spending a semester or an academic year overseas has already become an integral part of college life (IIE, 2016).

 Current Issues in Comparative Education 19(2), 111-139.

Given the growing proportion of undergraduate students with overseas learning experience, study abroad has become a significant part of the US higher education system that cannot be ignored by education researchers and policy makers. A key question for study abroad stakeholders is: what is the value-added of study abroad to college education?

China, a country that has traditionally sent students abroad, has started to involve more college students in study abroad programs that are sponsored by or affiliated with Chinese institutions, instead of letting Chinese students completely leave their Chinese institutions and enroll in a program at a foreign country as an international student. The most recent Chinese College Student Survey (CCSS) shows that 6% of college students studied abroad before graduation and approximately 35% of students are planning to do so.

Given the rapid increase of students studying abroad in higher education and the growing consensus that studying abroad provides some of the richest and most powerful forms of experiential learning for our students (Burn, 1991; Hamir, 2011; Kuh, 1995; Kuh et al., 2008; Laubscher, 1994; McKeown, 2009; Tarrant et al., 2014; Li, 2016), policymakers and the general public have become increasingly interested in the potential impact of study abroad across disciplines. Extant literature on Chinese students studying abroad mostly focus on students who leave China completely to pursue an overseas degree and potentially cause brain drain. Very few prior-studies examine the behavior of Chinese college students studying abroad in non-degree programs. As a comparison, the main stream of American college students studying abroad is in non-degree programs. Thus, results from the US may provide a good reference.

Hence, the goal of this research is therefore to quantify the value of studying abroad during college using descriptive analysis and quasi-experimental methods such as Propensity Score Matching with a good outcome measure (e.g., academic achievement) that is available in both the US and China. In this research, we choose GPA as a proxy for academic achievement. The data used in this research comes from the National Survey of Student Engagement (NSSE) in the US and the CSSS in China, both collected in 2014. After investigating student profile and estimating the impact of studying abroad on college students' academic development in the US, and China separately, this research also conducts a comparison by examining the similarities and differences

between the two countries. This research is designed to address the following research questions:

> 1) In the U.S, who studied abroad during college and what are the impacts of study abroad on undergraduate students' academic achievement?
>
> 2) In China, who studied abroad during college and what are the impacts of study abroad on undergraduate students' academic achievement?
>
> 3) What are the similarities and differences between the US and China, in terms of study abroad student profile and the impact of this experience on students' academic achievement?

The remainder of this paper is organized as follows: Section Two reviews existing literature on the impact of studying abroad on student academic achievement; Section Three describes the data source; Section Four introduces the empirical strategies and sample; Section Five presents the results based on Ordinary Least Square regression and Propensity Score Matching, by using studying abroad as an indicator; and Section Six concludes the paper with a summary of findings, limitations, and further analysis.

Literature Review

Regarding the value of study abroad, a large body of study abroad research focused on intercultural understanding and reported a positive impact of study abroad on student global engagement, cross-cultural adaptability and cultural sensibility (Carlson, 1990; Carlson, 1998; Paige et al., 2002; Chieffo, 2004; Kauffmann et al., 1992; Kitsantas, 2004; Williams, 2005; Anderson et al, 2006; Black & Duhon, 2006; Lewin, 2010; Li, 2016). More recent studies looked into the relationship between study abroad and labor market outcomes such as employment and starting salary. Findings from these studies are consistent: study abroad has positive impact on student career path, earnings and employment (Paige, 2009; Palifka, 2009; Salisbury, 2009; Li, 2016). Trooboff (2008)'s results also shed light of the mechanism because employers value study abroad in hiring recent college and university graduates. However, the findings on academic achievement are mixed. A number of studies have evaluated the effects of study abroad experience on student academic development. Overall, evidence from

these studies was mixed, depending on the data, outcome measure, program type, and methodology. However, there are more studies reporting positive effects.

On the positive side, many studies (Allen, 2009; Barron, 2003; Diao, &Freed, 2011;, Engle, 2012; Foster, 2001; Freed, 1995; Freed, 1998; Freed, Segalowitz, & Dewey, 2004; Jimenez-Jimenez, 2010; and Segalowitz et al., 2004) reported students gained foreign language proficiency from their studies overseas. Engle and Engle (2004), Kinginger and Farrell (2004), Allen (2009), and Diao and Freed (2011) found positive evidence in French grammar and vocabulary of students by using a pre-/post-test comparison. Similar evidence was found in other language programs such as for Spanish (Jimenez-Jimenez, 2010; Segalowitz et al., 2004), Chinese (Foster, 2001), and Hebrew (Donitsa-Schmidt & Vadish, 2005). Regarding concerns that studying abroad may delay timely college completion, Xu et al., (2013) found beneficial effects of study abroad programs on undergraduate degree completion. O'Rear et al., (2014) also confirmed that studying abroad can increase the likelihood of college graduation. Li (2016) found that study abroad during college has positive impact on student GPA.

On the negative side, Wilkinson (1998) interviewed four students and challenged the idea that study abroad facilitates student language acquisition; Savicki et al., (2012) reported no evidence that studying abroad improves student language acquisition from the two programs to Austria and Spain; Mendelson's (2004) assessment did not find positive evidence of student academic achievement either.

It is important to bear in mind, however, that despite the results, sample sizes of these studies are commonly small and only one study identify a control group (Jimenez-Jimenez, 2010) and its selection criteria remains questionable. Even though small group studies may provide deep understanding of a behavior where large datasets can fail, Jimenez-Jimenez chose a group of six native speakers from Spanish-speaking countries as a control group to study abroad students who learned Spanish as a second language, which was not the best comparison for many reasons. The two undergraduate degree completion studies, Xu et al., (2013) and O'Rear et al., (2014), addressed the sample size problem by using data from one college (Old Dominion University) and from one state (Georgia), but one college and one state is hardly representative. Thus, solid research with national-level data is needed to check the external validity of these studies in order to obtain an accurate estimate of the real effect of study abroad on a student's academic achievement.

Even when national-level datasets are available, a comparative study might still not be feasible, unless the data from different countries are collected with the same (or similar enough) survey design and outcomes are measured in the same way. By the same token, even though studies reporting positive effects seem to be more in number than other studies, heterogeneous effects by sending countries (i.e., the impact of study abroad on American students are substantially different from that on Chinese students) remain unknown. There is far too little understanding of the theory and practice of this type of learning in the context of comparative and international education.

Data Source

This study uses two comparable national surveys in the US and China to estimate and compare the impact of study abroad programs on student academic achievement. The American dataset used is from the National Survey of Student Engagement (NSSE) 2014 conducted by Indiana University.[1] The Chinese dataset we used is the restricted use data from the Chinese College Student Survey (CCSS), which is the largest national survey on college students in China. Specifically, we used the Questionnaire for Undergraduate Students in Four-year Institutions, which was developed from the NSSE questionnaire through the collaboration between Indiana University and Tsinghua University. Based on our knowledge, the CCSS is the only national survey in China that contains information about study abroad during college. More importantly, because the questionnaire is developed from NSSE, the CCSS has many similarities with the NSSE. Given the substantial similarities[2] between these two surveys, this pair of datasets provides a unique opportunity to conduct a comparative study that perfectly fits this research's purpose.

[1] To the best of our knowledge, the NSSE, the Student Experience in the Research University Survey (SERU), and the College Senior Survey (CSS) are the only three US national college experience surveys that contain information on study abroad. These surveys focus on how students evaluate their experience at the higher education institutions they attend, and each survey has its own pros and cons. We chose the NSSE over SERU and CSS because the NSSE is larger and the most spread-out sample than the other two. For example, the NSSE 2014 contains data on 29,836 senior students from 622 US colleges and universities vs. 23,523 senior students from 95 US institutions in the CSS 2014. Another example from SERU is that the survey excludes non-research colleges and universities, such as liberal arts colleges, which are the majority institutions with respect to both undergraduate enrollment and study abroad. Therefore, the NSSE is the most desirable dataset for this research.

[2] Like NSSE, the CCSS examines student engagement at college with a particular focus on higher impact activities such as study abroad. And because the CCSS employs NSSE's survey design, CCSS's instruments, measurements and the coding structure are very similar to NSSE as well. For instance, both CCSS and NSSE ask students how many hours per week they spend to prepare their classes and specify that "preparing for class" include studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, and other academic activities.

Methodology

Quantitative identification of the causal effect of study abroad is difficult because of data constraints and problems such as endogeneity (i.e. participation in study abroad programs is determined by variables which also determine the dependent variable). This research uses a multivariate regression model and a Propensity Score Matching (PSM) strategy to address the endogeneity problem.

Baseline Model The baseline model to estimate the effect of studying abroad on student academic achievement is specified as following:

$$Y = \alpha + \beta Z + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \varepsilon \quad (1)$$

This model builds upon the standard education production function. Where Y refers to academic achievement, Z is a dummy variable indicating whether or not the student studied abroad during college, and X is the vector of confounding covariates, and ε is the error term that follows a normal distribution. More specifically, we use the self-reported Grade Point Average (GPA) to measure academic achievement. In the NSSE 2014 survey, students were asked to report the most grades received on courses. We convert grades to GPA on a scale of 1-4 based on the guideline of the College Board,[3] and then rescale the GPA into 1-100 scale to make it comparable with the Chinese data. The CCSS asked for the average course grade in "last semester" (1-100 scale). As the GPA measures are not directly comparable across countries, we estimate the models separately for the US and the Chinese samples, and compare the effect size of the estimates.[4] With regards to the covariates, we include variables that are correlated to academic performance as suggested by the literature (Freed, 2014; Salisbury, 2009). Specifically, X_1 stands for student characteristics and college experience, including age, gender, ethnicity, academic major, enrollment status (US only), whether live on campus (US only), and hours per week spent on course-related work, X_2 are the family covariates including parents' education level, whether live in urban area (China only), and whether the student is the only child in his/her family (China only), $X_3$are institutional characteristics including the type, size, and location of the institution. The variables are constructed with self-reported information in the

[3] http://www.collegeboard.com/html/academicTracker-howtoconvert.html

[4] GPA may not be comparable across universities. A conventional way to address this problem is to include institution fixed-effect in the model. However, when institution fixed-effects were included, some individual and family variables were omitted because of collinearity. Thus we decided to control for available institutional characteristics that may result in differences in grading.

NSSE 2014 and CCSS 2014. Tables A1 and A2 (see Appendix) report the measurement and descriptive statistics of the variables used for the US and the Chinese samples, respectively.

Propensity Score Matching In order for the Ordinary Least Square (OLS) estimate of the coefficient on study-abroad program to be the true causal effect, participation in study abroad programs must be exogenous to student academic performance. But we are concerned about the selection bias that could arise because students are not randomly assigned to such programs, but self-selected. It could be argued that students from affluent families are more likely to go abroad as they can obtain extra financial support from parents, which is implied by the descriptive analysis. Affluent families are also more likely to afford extra educational services that assist students with their coursework such as data/software licenses, private tutoring, and professional editing. Consequently, the between-group difference in academic achievement may not be caused by study abroad experience, but because the study-abroad group is ultimately different from the non-study-abroad group. In other words, these two groups are not comparable at all. If that is the case, estimates from the OLS regression will be biased.

In order to identify the true treatment effect, this study endeavors to address the endogeneity problem with Propensity Score Matching (PSM). Propensity score theory states that rather than controlling for all the variables, it is sufficient to control for just the propensity score, which is just a one-number summary of the covariates.

$$e(X) = \Pr(Z = 1 \mid X) \tag{2}$$

where Z is the treatment variable (participating study abroad in this research) and X is a vector of pre-treatment covariates. In practice, this means using the matched groups to estimate each mean. The primary advantage of Propensity Score Matching over Ordinary Least Square is that Propensity Score Matching does not have to specify the multi-dimensional relationship between X and the outcome. This way, Propensity Score Matching also reduces bias caused by possible multicollinearity among covariates. In this study, the estimand is Average Treatment effect on the Treated (ATT). In terms of estimation strategy, Probit regression is used to estimate this effect

size.[5] The results reported in this paper are from use of the nearest neighbor method without replacement.[6]

Analytic Sample The American data used is a 20% random sample from the NSSE 2014, which was collected by Indiana University and shared with Teachers College, Columbia University.[7] It contains 5,361 students from 71 universities and college. Institutions' participation in the NSSE survey is voluntary, and each institution uses different sampling strategies to select students. The response rate varies across institutions with an average of 31% in 2014. Although the NSSE sample cannot be considered as a strictly representative sample of the US undergraduate students without a random sampling scheme, its demographic composition is overall similar to the US Bachelor's-Granting institutions and students, with a little over-representative of public colleges and universities and master's colleges and universities under the Carnegie Basic Classification.

The Chinese sample from the CCSS 2014 survey contains 55,529 students from 38 four-year universities and colleges in China.[8] Similar to NSSE, the participation in the CCSS survey is voluntary. But in each participating institution, the student sample is selected with a stratified random sampling strategy. The overall response rate of the 2014 survey is 67%. The sample is a little over-representative of Project "985" and Project "211" universities (i.e. the two elite university projects in China) and universities specialized in science and engineering. In order to make the samples representative for both countries, institutional and individual sampling weights provided by the NSSE team and the CCSS teams are applied in all analyses.

Empirical Results

Descriptive analysis Overall, studying abroad during college is more common in the US than in China. The American representative sample used in this research includes more than 5,500 students who were attending college in 2014. Among them, approximately 14% of students have study abroad experience. The Chinese data from

[5] Probit regression is a conventional way to estimate models with binary dependent variable. The error term is assumed to follow the standard normal distribution. Probit models are most often estimated with the standard maximum likelihood procedure.

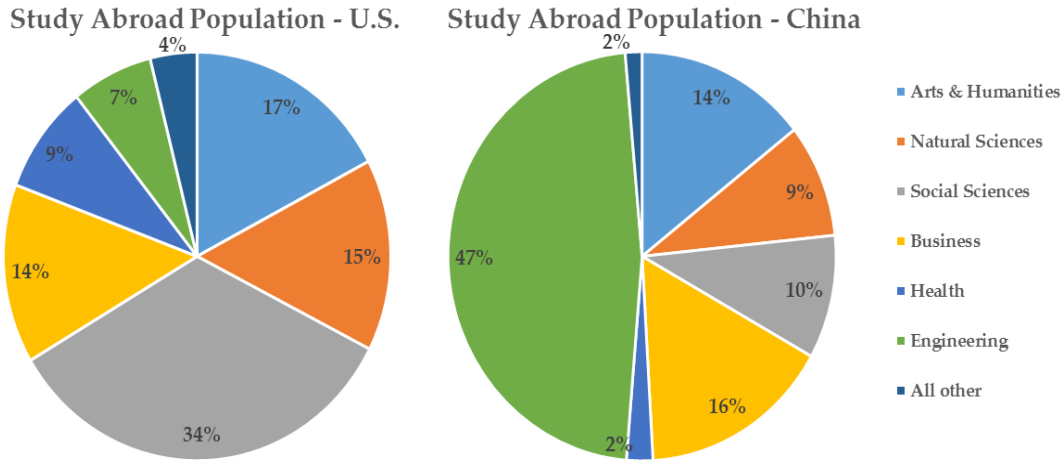
[6] The nearest neighborhood matching is one of the methods to construct comparable control group for the treatment group. It chooses the control unit that is closest to the treated unit on a given distance measure based on the propensity score. The matching is done one at a time for each individual in the treatment group. Once a match is selected, it will not be replaced back to the untreated pool.

[7] According to NSSE's data sharing policy, only a random sample up to 20% of the full sample is available.

[8] We were able to have access to the full sample of CCSS 2014.

over 55,500 college students in the same year shows a 6% participating rate in study abroad, less than half of the American participating rate.

Figure 1. Disciplinary Distribution of Study Abroad Population, by field of study



The composition of study abroad population is also different. Female students are underrepresented (43% in the study abroad sample vs. 47% in the full sample) in China and overrepresented in the US (75% female in the study abroad sample vs. 65% female in the full sample). Figure 1 breaks down the study abroad populations by their field of study. Engineering program accounts for 47% of the Chinese study abroad population but only 7% in the U.S.[9]

With regard to the differences in study abroad behavior by sub groups, Table 1 presents the percentage of students that studied abroad in each groups, which can be interpreted as how likely a student from that group will study abroad. In general, female students are more likely to study abroad in the US but less likely in China. In both countries, students in Arts & Humanities, Social Sciences, and Business majors are more likely to study abroad than other majors. In contrast, the percentage of Health students studied abroad is dramatically low.

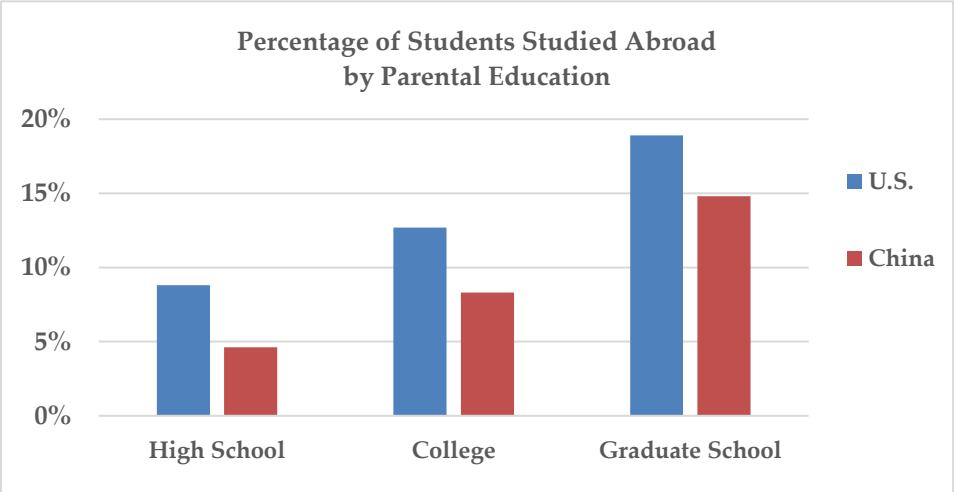
[9] This could be explained by sampling difference. The Chinese sample is over-representative of engineering students. According to the national statistics of 2014, engineering students accounted for 33% of the total enrollment of four-year undergraduate students. The percentage in CCSS 2014 is 46.6% after weights are applied. This is because the sample contains a higher percentage of engineering concentrated universities and colleges.

Table 1. Percentage of Students Studied Abroad

	US		China	
	Number of Students	% of students studied abroad	Number of Students	% of students studied abroad
Total	5,631	13.9	55,529	6.1
Gender				
Male	1,968	10.3	28,592	6.6
Female	3,663	15.9	26,937	5.7
Major				
Arts & Humanities	615	21.8	6,263	7.8
Natural Sciences	906	13.1	5,877	5.7
Social Sciences	1,642	16.3	4,304	7.1
Business	850	13.1	7,750	7.0
Health	792	8.5	3,163	2.3
Engineering	452	11.9	27,041	6.0
All other	374	16.7	1,131	4.1
Enrollment Size				
Small	1,802	15.1	9,904	5.9
Medium	1,236	14.6	28,937	6.6
Large	2,453	13.3	16,688	5.5
Parental Education				
High School	1,067	8.8	36,444	4.6
College	2,773	12.7	15,878	8.3
Graduate School	1,791	18.9	1,878	14.8

Parental education turns out to be a strong indicator of students' study abroad behavior. As indicated in Figure 2, both in the US and China, students with highly educated parents are more likely to study abroad compared to other students. The US-China gap is quite small among students with parents with a graduate school education. One possible explanation could be that highly educated parents are more likely to have a high income to pay for their children's study abroad programs. Another possibility could be that highly educated parents have international exposure themselves. Compared to less educated parents, these parents tend to be more open-minded and willing to support their children studying abroad. Having an urban background is another strong determinant to studying abroad behavior. In particular, the participating rate of Chinese students from rural areas is remarkably low. Again, family's socio-economic status determines students' decisions about studying abroad.

Figure 2. Percentage of Students Studying Abroad by Parental Education



Multivariate regression The baseline Ordinary Least Square regression model based on Equation (1) is established to investigate the average effect of study abroad. We ran three sets of regressions on each dataset and report the results in Table 2. Models 1 and 4 include student demographic characteristics; Models 2 and 5 include student and family variables; Model 3 and 6 are the full specification by adding institutional control variables.

The estimations across these three models consistently report a statistically positive impact of study abroad on student academic achievement, in both the US and China settings. Results indicate about a roughly three point increase[10] in GPA on a scale of 1-100 for American students and one-point increase[11] for Chinese students. Results also suggest the following: 1) in both the US and China, female students outperform male students; compared to students in Arts and Humanities, students in other majors tend to have a lower GPA; majority students (white American students and Han Chinese students) outperformed minority students; 2) in terms of family background, parental education is not associated with student GPA in the US but a strong predictor of Chinese students' academic achievement. In China, the father's education turns out to be positively associated with student GPA while no such evidence is found in the relationship between mother's education and student academic outcomes, after controlling for institutional characteristics; and 3) students from American private institutions or Chinese elite colleges and universities (e.g., Chinese Project 985 and

[10] Effect size is 0.22 standard deviation for American students
[11] Effect size is 0.10 standard deviation for Chinese students

Project 211 higher education institutions that receive exclusive funding from the central government) outperformed students of other institution types.

Table 2. OLS Regression Results

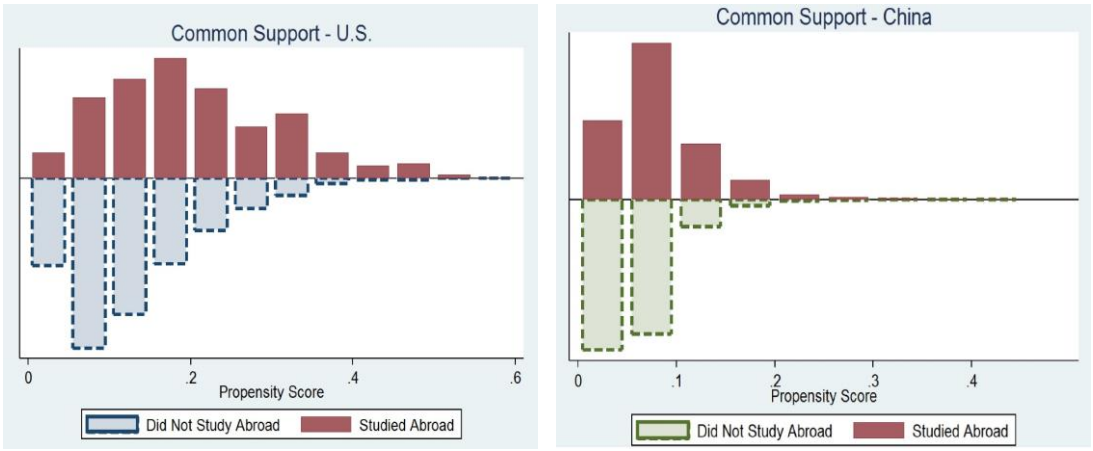
US				China			
Variable	Model 1	Model 2	Model 3	Variable	Model 4	Model 5	Model 6
abroad	3.02***	2.84***	2.77***	abroad	0.96***	0.91***	0.97***
	(0.44)	(0.46)	(0.46)		(0.21)	(0.21)	(0.20)
hours	0.22***	0.21***	0.21***	hours	0.17***	0.17***	0.14***
	(0.02)	(0.02)	(0.02)		(0.01)	(0.01)	(0.01)
age	0.18***	0.21***	0.20***	age	-0.04***	-0.04***	-0.03***
	(0.02)	(0.03)	(0.03)		(0.01)	(0.01)	(0.01)
male	-1.17***	-1.22***	-1.20***	male	-3.16***	-3.15***	-3.29***
	(0.38)	(0.39)	(0.39)		(0.11)	(0.11)	(0.10)
white	2.48***	2.53***	2.59***	minority	-1.20***	-1.26***	-1.97***
	(0.54)	(0.55)	(0.56)		(0.17)	(0.17)	(0.17)
black	-5.40***	-5.07***	-5.12***	party	4.69***	4.66***	4.33***
	(0.89)	(0.90)	(0.91)		(0.15)	(0.15)	(0.15)
hispanic	-1.17	-0.47	-0.70	m_social	-0.55	-0.59	-0.77
	(0.80)	(0.83)	(0.85)		(0.23)	(0.23)	(0.22)
asian	1.79*	2.06**	1.82*	m_natural	-1.49***	-1.53***	-1.99***
	(0.94)	(0.93)	(0.95)		(0.21)	(0.21)	(0.21)
m_social	-1.80***	-1.62***	-1.51***	m_business	-1.30***	-1.34***	-1.11***
	(0.57)	(0.57)	(0.57)		(0.20)	(0.20)	(0.19)
m_natural	-2.31***	-2.31***	-2.18***	m_health	-1.95***	-2.03***	-2.20***
	(0.65)	(0.66)	(0.66)		(0.25)	(0.25)	(0.27)
m_business	-2.67***	-2.59***	-2.64***	m_engineer	-1.48***	-1.52***	-1.56***
	(0.66)	(0.66)	(0.66)		(0.17)	(0.17)	(0.17)
m_health	-2.91***	-2.69***	-2.73***	m_other	-2.02***	-2.06***	-2.43***
	(0.68)	(0.69)	(0.69)		(0.37)	(0.37)	(0.37)
m_engineer	-5.49***	-5.58***	-5.34***	onlychild		-0.36***	-0.33***
	(0.86)	(0.86)	(0.88)			(0.11)	(0.11)
m_other	-4.43***	-4.20***	-4.26***	urban		-0.46***	-0.64***
	(0.88)	(0.88)	(0.88)			(0.11)	(0.11)
				Father			
Full-time		-0.03	-0.03	College+		0.81***	0.37**
		(0.66)	(0.66)			(0.18)	(0.17)
				Mother			
Transfer		0.39	0.36	College+		0.45**	0.21
		(0.49)	(0.49)			(0.20)	(0.20)
dorm		0.45	0.38	medium			-0.59***
		(0.41)	(0.42)				(0.17)
first		-1.77***	-1.75***	large			-1.61***
		(0.52)	(0.52)				(0.19)
pedu_cl		0.75	0.80	inst_985			5.43***
		(0.59)	(0.59)				(0.19)

pedu_gs		0.69	0.73	inst_211			2.07***
		(0.74)	(0.74)				(0.17)
private			1.50***	inst_univ			0.57***
			(0.56)				(0.15)
medium			0.87	munici			-1.05***
			(0.55)				(0.12)
large			0.96				
			(0.64)				
r_neweng			1.01				
			(0.70)				
_cons	79.51***	78.58***	77.40***	_cons	79.13***	79.53***	79.73***
	(0.92)	(1.50)	(1.59)		(0.31)	(0.32)	(0.34)
N	5,477	5,387	5,387	N	36,736	36,736	36,736
R-Squared	0.08	0.08	0.08	R-Squared	0.08	0.09	0.12

Note: a) Coefficients are reported; b) Robust standard errors are in parentheses; c) Reference group for race, major, region, and enrollment size are other race (US data) or minority (China data), Arts and Humanities majors, and small enrollment size (American institutions with under 2,500 enrollment or Chinese institutions with under 20,000 enrollment); and d) * p<0.1; ** p<0.05; *** p<0.01.

Propensity score matching To address selection bias, we employ a Propensity Score Matching approach. We first compute propensity scores for each student, then find a match for each study-abroad student from the non-study-abroad group based upon their propensity score, run a weighted regression on the matched sample, and estimate the Average Effect on the Treated (ATT). Because Propensity Score Matching is only appropriate when certain assumptions hold, such as common support and balance, we check the common support in Figure 3.

Figure 3. Propensity Score Matching – Common Support, by country



Specifically, we plot the distribution of propensity scores with the study-abroad group in the top histogram and the non-study-abroad in the bottom histogram. The

horizontal axis indicates propensity scores and the length of each bar on the vertical axis indicates the fraction of sample falling into a corresponding interval of propensity score. As we can see from Figure 3, the overlap of the propensity score is pronounced, especially in the lower half. Only a few observations are off support at the high end of the propensity score.

The critical task for Propensity Score Matching is to find a balanced model. This study uses a rule of thumb to define "good balance": 1) for continuous variables, the difference in means is less than or equal to .05; treatment group (study abroad students) standard deviations and the ratio of standard deviations are between .91 and 1.1; and 2) for binary/indicator variables, the difference in percentage across groups is less than or equal to .025. Table A3 (see appendix) summarized the balance check for each of the variables. Overall this model satisfies the balance criteria except the variable *age*, not perfectly but as well as we can achieve. Regression results from Propensity Score Matching are reported in Table 3, along with heterogeneous effects.

Heterogeneous effects and robustness check For heterogeneous effects, we seek to determine how individual characteristics interact with the study abroad experience in terms of academic achievement. Hence, we examine how the impact of study abroad varies depending on individual characteristics, such as gender, race, and major. These potential differences are tested by adding a series of interaction terms to the final regression model (Model 3) on the matched sample. The interaction is generated as the product of two dummy variables. Study abroad is coded as a dummy, as are student characteristics. For example, the *AbroadXmale* variable is the interaction of study abroad and gender, which is computed by the study abroad variable (1=studied abroad; 0=did not study abroad) multiplied by the gender variable (1=male; 0=female). The coefficient of *AbroadXmale* indicates the gender difference in the impact of studying abroad on student academic achievement. Table 3 summarizes the regression results for each of the above characteristics by adding one interaction at a time, along with the original results from Propensity Score Matching without interactions.

Table 3. Propensity Score Matching Results and Heterogenous Effects

US					China				
Variable	**PSM**	**Male**	**White**	**Natural Science**	**Variable**	**PSM**	**Male**	**Han**	**Natural Science**
abroadRXmale		2.61*			abroadXmale		0.45		
		(1.60)					(0.77)		
abroadRXwh~e			0.84		abroadXhan			0.67	
			(1.61)					(1.16)	
abroadRXm_natural				-0.19	abroadXm_natural				-0.19
				(0.81)					(0.78)
abroadR	2.86***	2.18***	2.24***	2.95***	abroad	0.75*	0.50	0.13	0.86
	(0.74)	(0.89)	(1.35)	(0.83)		(0.39)	(0.55)	(1.09)	(0.59)
hours	0.22***	0.22***	0.22***	0.22***	hours	0.11***	0.11***	0.11***	0.11***
	(0.04)	(0.04)	(0.04)	(0.04)		(0.02)	(0.02)	0.02	0.02
age	0.13	0.13	0.13	0.13	age	0.00	0.00	0.00	0.00
	(0.09)	(0.09)	(0.09)	(0.09)		(0.05)	(0.05)	(0.05)	(0.05)
male	-0.89	-2.18	-0.89	-0.89	male	-2.94***	-3.16***	-2.94***	-2.94***
	(0.81)	(1.34)	(0.81)	(0.81)		(0.40)	(0.58)	(0.40)	(0.40)
white	0.48	0.46	0.05	0.48	han	1.98***	1.97***	1.62***	1.99***
	(1.14)	(1.14)	(1.56)	(1.13)		(0.61)	(0.61)	(0.75)	(0.61)
black	-11.06***	-11.03***	-11.03***	-11.07***	party	3.52	3.53	3.53	3.52
	(2.15)	(2.16)	(2.15)	(2.15)		(0.50)	(0.50)	(0.50)	(0.50)
hispanic	-1.35	-1.18	-1.36	-1.34	m_social	-1.03	-1.04	-1.03	-1.03
	(1.62)	(1.63)	(1.62)	(1.62)		(0.81)	(0.81)	(0.81)	(0.81)
asian	-4.73***	-4.86***	-4.74***	-4.74***	m_natural	-1.77**	-1.77**	-1.77**	-1.67*
	(1.91)	(1.92)	(1.91)	(1.92)		(0.84)	(0.84)	(0.84)	(0.99)
m_social	-1.92*	-1.90*	-1.91*	-1.92*	m_business	-1.52**	-1.53**	-1.52**	-1.52**
	(1.05)	(1.05)	(1.04)	(1.04)		(0.75)	(0.75)	(0.75)	(0.75)
m_natural	-1.82	-1.76	-1.82	-1.22*	**m_health**	-1.68	-1.68	-1.67	-1.68

	(1.22)	(1.22)	(1.22)	(2.90)		(1.34)	(1.33)	(1.33)	(1.34)
m_business	-3.12***	-3.10***	-3.11***	-3.11***	m_engineer	-1.01	-1.01	-1.01	-0.91
	(1.15)	(1.15)	(1.15)	(1.15)		(0.64)	(0.64)	(0.64)	(0.81)
m_health	-3.22**	-3.21**	-3.24**	-3.21**	m_other	-4.37***	-4.37***	-4.38***	-4.36***
	(1.46)	(1.45)	(1.47)	(1.46)		(1.78)	(1.78)	(1.78)	(1.78)
m_engineer	-5.63***	-5.54***	-5.62***	-5.05***	onlychild	-0.28	-0.28	-0.29	-0.28
	(1.71)	(1.71)	(1.71)	(3.09)		(0.47)	(0.47)	(0.47)	(0.47)
m_other	-2.90	-2.68	-2.88	-2.89	urban	-0.99	-0.98	-0.99	-1.00
	(1.89)	(1.90)	(1.89)	(1.89)		(0.49)	(0.49)	(0.49)	(0.49)
ft	1.60	1.67	1.64	1.61	fabaplus	0.08	0.08	0.08	0.08
	(1.34)	(1.34)	(1.34)	(1.34)		(0.56)	(0.56)	(0.56)	(0.56)
trans	-0.67	-0.69	-0.68	-0.67	mobaplus	0.28	0.28	0.29	0.29
	(1.12)	(1.12)	(1.12)	(1.13)		(0.63)	(0.63)	(0.63)	(0.63)
dorm	-0.43	-0.44	-0.43	-0.43	large	-2.25***	-2.25***	-2.25***	-2.25***
	(0.85)	（0.86）	(0.85)	(0.85)		(0.80)	(0.80)	(0.80)	(0.80)
first	-0.19	-0.16	-0.16	-0.20	medium	-0.75	-0.76	-0.75	-0.75
	(1.14)	(1.14)	(1.14)	(1.14)		(0.62)	(0.62)	(0.62)	(0.62)
pedu_cl	2.35*	2.39*	2.37*	2.36*	inst_985	5.79***	5.79***	5.79***	5.79***
	(1.37)	(1.37)	(1.37)	(1.37)		(0.68)	(0.68)	(0.68)	(0.68)
pedu_gs	2.06	2.13	2.08	2.06	inst_211	1.97***	1.97***	1.97***	1.97***
	(1.66)	(1.67)	(1.66)	(1.66)		(0.64)	(0.64)	(0.64)	(0.64)
pri	2.31**	2.28**	2.31**	2.31**	inst_univ	0.28	0.28	0.28	0.28
	(1.07)	(1.07)	(1.07)	(1.07)		(0.58)	(0.58)	(0.58)	(0.58)
medium	1.54	1.55	1.54	1.54	munici	-1.58***	-1.58***	-1.58***	-1.58***
	(1.08)	(1.08)	(1.08)	(1.08)		(0.49)	(0.49)	(0.49)	(0.49)
large	4.39***	4.35***	4.38***	4.39***					
	(1.33)	(1.33)	(1.32)	(1.33)					
r_neweng	1.53	1.49	1.50	1.53					
	(1.44)	(1.46)	(1.44)	(1.45)					

_cons	76.04***	76.34***	76.32***	75.95***	_cons	80.38***	80.50***	80.37***	80.33***
	(3.62)	(3.64)	(3.70)	(3.62)		(1.43)	(1.48)	(1.43)	(1.47)
N	1,280	1,280	1,280	1,280	N	3,568	3,568	3,568	3,568
R-Squared	0.11	0.11	0.11	0.11	R-Squared	0.09	0.09	0.09	0.09

Note: a) Coefficients are reported; b) Robust standard errors are in parentheses; c) Reference group for race, major, region, and enrollment size are other race (US data) or minority (China data), Arts and Humanities majors, and small enrollment size (American institutions with under 2,500 enrollment or Chinese institutions with under 20,000 enrollment); and d) * p<0.1; ** p<0.05; *** p<0.01.

Table 4. Corrected Standard Errors with Bootstrap (US) - Sample Size: 5,453

Variable	Replication	Observed	Bias	Std. Err	[95 % Conf. Interval]	
_bs_1	1,000	2.60	-0.07	0.87	0.88	4.32 (N)
					0.80	4.33 (P)
					0.94	4.51 (BC)

Note: N=normal; P=percentile; BC=bias-corrected

Table 5. Corrected Standard Errors with Bootstrap (China) - Sample Size: 38,634

Variable	Replication	Observed	Bias	Std. Err	[95 % Conf. Interval]	
_bs_1	1,000	1.13	-0.04	0.42	0.31	1.94 (N)
					0.34	1.98 (P)
					0.43	2.13 (BC)

Note: N=normal; P=percentile; BC=bias-corrected

Three main observations emerge from Table 3. First, the coefficient of studying abroad in the US sample is consistently significant across models, which indicates that study abroad experience has a robust positive impact on student academic achievement across gender, race, and major. Second, the coefficient of studying abroad turns out to be insignificant after adding gender. This suggests that for Chinese students, the impact of study abroad on student academic achievement is not as robust as for American students. Third, the gender interaction term is significant in the US sample. This means that being male increases the impact of study abroad on student academic achievement. In other words, male American students benefit more from study abroad than female American students. There are no such gender differences among Chinese students participating in study abroad.

We use bootstrap to simulate the matching procedure 1,000 times to get the distribution of the average treatment on the treated (ATT)[12]. As reported in Table 4 and Table 5, on average, for students who studied abroad, having that experience increased GPA by 2.60 points for American students and 1.13 points for Chinese students. These simulation results are consistent with the finding from the baseline regression, Propensity Score Matching, and heterogeneous effects. Overall, on average, study abroad experience at college seems to have a larger and more consistent impact on American student academic achievement than on Chinese students.

Conclusions and Discussion

Overall, we find significant positive impact of study abroad on student academic achievement in both countries, with higher impact for American students. Results from the NSSE data reveal a positive impact of study abroad on students' academic achievement during college. On average, study abroad experiences improve students' GPA by 3 points (on a scale of 1 to 100). For comparison, China's study abroad group is more evenly distributed across gender and the field of study. However, descriptive analysis reveals that students from rural areas are less likely to study abroad. Propensity score matching results indicate a slight (one point) positive impact on students' academic achievement in China. These results are consistent with Freed (2004) and Ingraham, & Peterson (2004)'s findings that study abroad during college has a positive impact on GPA and facilitate students' intellectual growth. This increase in GPA may translate into benefits for the labor market, as research also shows that

[12] Bootstrap is a resampling method to get robust estimates on parameters through estimation of sampling variability. We resample the original sample for 1,000 times and conduct propensity score matching on each redrawn sample to estimate the distribution of average treatment on treated (ATT) effect.

students with high GPA have an advantage in the job market because of human capital and signaling effects (Jones, & Jackson, 1990).

Regarding studying abroad student profiles, the majority of American students who study abroad are still comprised of the traditional study-abroad student body: white female in Arts and Humanities or Social Science programs at a private institution. The gender and major differences in Chinese students are smaller than American students. Importantly, parental education turns out to be the strongest predictor of students' study abroad behavior. Students with highly educated parents are more likely to study abroad compared to other students. The US-China gap is quite small among students with graduate-school-educated parents. Urbanicity is a strong indicator as well. In particular, Chinese students from large cities (such as direct-controlled municipality) are much more likely to join study abroad programs than students from other places. These finding confirms Goldstein, & Kim (2006) and Salisbury, et al., (2009)'s arguments that social economic status (SES) matters for American students and proves that the same for Chinese students.

Thus, aids to students from low SES families are solely needed to improve education equality both in the US and China. From a policy perspective, engaging college students in non-degree study abroad programs without completely leaving their Chinese institution will alleviate concerns about potential brain drain (Bhagwati, & Hamada, 1974; Mountford, 1997; Beine, Docquier, & Rapoport, 2001; Beine, Docquier, & Rapoport, 2008; Mountford, & Rapoport, 2011).

This research contributes to the extant literature in several ways: a) it presents the comparison and contrast between two countries (instead of within one country); b) it employs quasi-experimental methods to estimate the effect size of study abroad experience on students (adding onto descriptive & correlation analysis); and c) it utilizes data from two national surveys that follow almost identical survey designs, which is rare in the field. Yet, we need to be cautious that our findings are subject to social and economic differences between the US and China.

Nonetheless, there are a few limitations of this study. First, the study abroad measure is an indicator that only captures whether or not students studied abroad, but does not differentiate by duration, subject, and organization of the study abroad program students attended. Dwyer (2004) and Carlson, & Widaman (1990) found long-term effects and McKeown (2009) reported the first-time effect. Hence, further comparative studies can exploit the variation of study abroad programs and evaluate effects of

various program design (e.g., duration, location, and curriculum) on student outcomes. Li (2016) conducted research on American students but there is no such study on Chinese students.

Second, this research focuses only on academic achievement. As Lewin (2010) pointed out that students benefit from study abroad through three ways: language study, courses in the student's major that offers new perspectives and knowledge acquisition. Future research can measure non-academic outcomes such as global mindset and labor market outcomes. Since 2013, NSSE incorporated a module of Global Perspectives to measure student cross-cultural awareness and attitude (Li, 2016). A new comparative study will be feasible if CCSS includes Global Perspectives module in the Chinese survey.

Finally, in order to truly understand the mechanism of how study abroad during college affects student outcome, we need to be able to track students (both study abroad alumni and their peers) during and after their college life. Qualitative research (i.e., interviews, focus group) will also provide deep understandings that large survey fails to capture.

Authors' Acknowledgement: The authors of this research thank the Center on Chinese Education at Teachers College, Columbia University, National Survey on Student Engagement (NSSE) at Indiana University, College Student Study Survey (CSSS) team at Tsinghua University for their data and support.

About the Authors:

Jing Li is a research associate of Economics and Education at Teachers College, Columbia University, where she received her Master's and PhD degrees, as well as rigorous training on both quantitative and qualitative research methods. Her dissertation is a quantitative research in which she employed economic theory and quasi-experimental methods to evaluate the impact of studying abroad experience on student outcomes. Her primary areas of study are labor economics, higher education policy and comparative & international education, with a focus on cross-cultural intervention, financial aid, racial differential and programmatic barriers to college success and career development in the global economy.

Fei Guo is a post-doctoral research fellow at the Institute of Education in Tsinghua University, Beijing, China. She holds an MA and a PhD in Economics and Education from Teachers College, Columbia University. She received a BA in International Economics from Shanghai Jiao Tong University in China. Fei Guo conducts research to explore factors influencing college students' academic achievement, cognitive and non-cognitive development, and labor market performance. Her research interests include student learning and development in higher education, economic returns to higher education, and educational equality and equity in China.

You You is a researcher at Institute of Economics of Education (IEE)/Graduate School of Education (GSE), Peking University. She received her PhD in Economics and Education from Teachers College, Columbia University in the City of New York. Her research interests include teachers labor market, college graduates employment and entrepreneurship, and education finance and she has expertise in the use of quantitative methods to evaluate the effects of educational policies and programs. She publishes widely and her research has won her several external research funds from national and provincial governmental agencies (e.g., National Natural Science Foundation of China).

***Contact**:* Jing Li at *jing.li@tc.columbia.edu*, Fei Guo at *feiguo0121@mail.tsinghua.edu.cn*, and You You at *yyou@pku.edu.cn*

References

Anderson, P., Lawton, L., Rexeisen, R., & Hubbard, A. (2006). Short-term study abroad and intercultural sensitivity: A pilot study. *International Journal of Intercultural Relations*, *30*(4), 457–469.

Allen, H. (2010). Interactive contact as linguistic affordance during short-term study abroad: Myth or reality? *Frontiers: The Interdisciplinary Journal of Study Abroad*, *19*, 1–26.

Barron, A. (2003). *Acquisition in interlanguage pragmatics: Learning how to do things with words in a study abroad context*. John Benjamins Publishing.

Beine, M., Docquier, F., & Rapoport, H. (2001). Brain drain and economic growth: Theory and evidence. *Journal of Development Economics*, *64*(1), 275–289.

Beine, M., Docquier, F., & Rapoport, H. (2008). Brain drain and human capital formation in developing countries: Winners and Losers*. *The Economic Journal*, *118*(528), 631–652.

Bhagwati, J., & Hamada, K. (1974). The brain drain, international integration of markets for professional and unemployment: A theoretical analysis. *Journal of Development Economics*, *1*(1), 19–42.

Black, H. T., & Duhon, D. L. (2006). Assessing the impact of business study abroad programs on cultural awareness and personal development. *Journal of Education for Business*, *81*(3), 140–144.

Burn, B. B. (1991). *Integrating study abroad into the undergraduate liberal arts curriculum: Eight institutional case studies*. New York, NY: Praeger.

Carlson, J. S. (1990). *Study abroad: The experience of American undergraduates*. Greenwood Publishing Group.

Carlson, J., & Widaman, K. (1988). The effects of study abroad during college on attitudes toward other cultures. *International Journal of Intercultural Relations*, *12*(1), 1–17.

Chieffo, L. (2004). Large-scale assessment of student attitudes after a short-term study abroad program. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 165–177.

Diao, W. (2011). Confirmed beliefs or false assumptions? A study of home stay experiences in the french study abroad context. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *21*, 109–142.

Donitsa-Schmidt, S. (2005). North american students in israel: An evaluation of a study abroad experience. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *11*, 33–56.

Dwyer, M. M. (2004). More Is Better: The Impact of Study Abroad Program Duration. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 151–163.

Engle, L. (2004). Assessing language acquisition and intercultural sensitivity development in relation to study abroad program design. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 219–236.

Engle, L. (2012). The rewards of qualitative assessment appropriate to study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *22*, 111–126.

Foster, P. (2001). A language and cultural practicum course in nanjing: Maximizing the students' use of chinese. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *7*, 121–128.

Freed, B. F. (1995). *Second language acquisition in a study abroad context*. John Benjamins Publishing.

Freed, B. F. (1998). An overview of issues and research in language learning in a study abroad setting. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *4*(2), 31–60.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, *26*(2), 275–301.

Goldstein, S. B., & Kim, R. I. (2006). Predictors of US college students' participation in study abroad programs: A longitudinal study. *International Journal of Intercultural Relations*, *30*(4), 507–521.

Hamir, H. (2011). *Go abroad and graduate on-time: Study abroad participation, degree completion, and time-to-degree.* Ph.D. dissertation, The University of Nebraska.

IIE. (2016). *IIE releases open doors 2016 data*. New York: Institute of International Education

Ingraham, E. C. , & Peterson. (2004). Assessing the impact of study abroad on student learning at michigan state university. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 83–100.

Jimenez-Jimenez, A. F. (2010). A comparative study on second language vocabulary development: Study abroad vs classroom settings. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *19*, 105–123.

Jones, E. B., & Jackson, J. D. (1990). College grades and labor market rewards. *The Journal of Human Resources*, *25*(2), 253–266.

Kauffmann, N. L., Marint, J. N., Weaver, H.D., & Weaver, J. (1992). *Students abroad: Strangers at home. Education for a Global Society.* Intercultural Press: Yarmouth, ME

Kinginger, C. (2004). Assessing development of meta-pragmatic awareness in study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 19–42.

Kitsantas, A. (2004). Studying abroad: The role of college students' goals on the development of cross-cultural skills and global understanding. *College Student Journal*, *38*(3), 441–452.

Kuh, G. D. (1995). The other curriculum: Out-of-class experiences associated with student learning and personal development. *The Journal of Higher Education*, *66*(2), 123–155.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *The Journal of Higher Education*, *79*(5), 540–563.

Laubscher, M. R. (1994). *Encounters with difference: Student perceptions of the role of out-of-class experiences in education abroad.* Westport, CT: Greenwood Press.

Lewin, R. (2010). *The handbook of practice and research in study abroad: Higher education and the quest for global citizenship*. Routledge.

Li, J. (2016). *The impact of study abroad on us undergraduate students' academic development, global perspective and labor market outcomes* (PhD Dissertation). Columbia University, New York.

McKeown, J. S. (2009). *The first time effect: The impact of study abroad on college student intellectual development*. Albany, NY: SUNY Press.

Mendelson, V. G. (2004). "Hindsight is 20/20:" Student perceptions of language learning and the study abroad experience. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 43–63.

Mountford, A. (1997). Can a brain drain be good for growth in the source economy? *Journal of Development Economics*, *53*(2), 287–303.

Mountford, A., & Rapoport, H. (2011). The brain drain and the world distribution of income. *Journal of Development Economics*, *95*(1), 4–17.

O'Rear, I., Sutton, R., & Rubin, D. (2014). *The effect of study abroad on college completion in a state university system*. Athens, GA: University of Georgia.

Opper, S. O. (1990). *Impacts of study abroad programmes on students and graduates. Higher Education Policy Series 11, Volume 2.* Jessica Kingsley Publishers: London

Paige, R. M., Fry, G. W., Stallman, E. M., Josić, J., & Jon, J.-E. (2009). Study abroad for global engagement: The long-term impact of mobility experiences. *Intercultural Education*, *20*(sup1), S29–S44.

Palifka, B. (2009). *Effects of study abroad and personality on employment and earnings*. VDM Publishing.

Savicki, V., Arrue, C., & Binder, F. (2012). Language fluency and study abroad adaptation. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *22*, 37–57.

Segalowitz, N. (2004). A comparison of spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*, 1–18.

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning spanish in at home and study abroad contexts. Studies in Second Language Acquisition, 26(2), 173–199.

Salisbury, M., Umbach, P., Paulsen, M., & Pascarella, E. (2009). Going global: Understanding the choice process of the intent to study abroad. *Research in Higher Education*, *50*(2), 119–143.

Tarrant, M. A., Rubin, D. L., & Stoner, L. (2014). The added value of study abroad fostering a global citizenry. *Journal of Studies in International Education*, *18*(2), 141–161.

Trooboff, S. B. (2008). Employer attitudes toward study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *15*, 17–33.

Wilkinson, S. (1998). Study abroad from the participants' perspective: A challenge to common beliefs. *Foreign Language Annals*, *31*(1), 23–39.

Williams, T. R. (2005). Exploring the impact of study abroad on students' intercultural communication skills: Adaptability and sensitivity. *Journal of Studies in International Education*, *9*(4), 356–371.

Xu, M., De Silva, C. R., Neufeldt, E., & Dane, J. H. (2013). The impact of study abroad on academic success: An analysis of first-time students entering Old Dominion University, Virginia, 2000-2004. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *13*.

Appendix

Table A1. Summary of Variables (US Dataset)

Variable	Observation	Mean	**Standard Deviation** (% for dummy variables)	Description
gpa100	5,613	85.67	13.44	GPA on a scale of 1 to 100 to be comparable with Chinese data
abroadR	5,567	0.14	0.35	Indicator of having studied abroad during college
hours	5,586	15.02	8.75	Hours per week spent on course-related work
age	5,590	23.09	7.82	Age at 2014
male	5,631	0.35	0.48	Male
white	5,631	0.66	0.47	Race: Caucasian
black	5,631	0.08	0.27	Race: African American
hispanic	5,631	0.07	0.25	Race: Hispanic
asian	5,631	0.05	0.22	Race: Asian
m_social	5,631	0.29	0.45	Major: Social Sciences
m_natural	5,631	0.16	0.37	Major: Natural Sciences
m_business	5,631	0.15	0.36	Major: Business
m_health	5,631	0.14	0.35	Major: Health
m_engineer	5,631	0.08	0.27	Major: Engineering
m_other	5,631	0.07	0.25	Major: Other major but not Education
ft	5,587	0.89	0.31	Full-time student
trans	5,605	0.26	0.44	Transferred student
dorm	5,605	0.40	0.49	Live on campus
first	5,609	0.39	0.49	First-generation attending college
pedu_cl	5,631	0.49	0.50	Parental education: College
pedu_gs	5,631	0.32	0.47	Parental education: Graduate Schools
pri	5,631	0.42	0.49	Institution control: private
medium	5,631	0.22	0.41	Institution size: medium
large	5,631	0.44	0.50	Institution size: large
r_neweng	5,631	0.05	0.23	Institution location: New England area

Table A2. Summary of Variables (China Dataset)

Variable	Observation	Mean	Standard Deviation (% for dummy variables)	Description
gpa	38,634	76.68	9.65	GPA on a scale of 1 to 100
abroad	55,529	0.06	0.24	Indicator of having studied abroad during college
hours	52,590	15.82	7.72	Hours per week spent on course-related work
age	55,529	22.61	7.97	Age at 2014
male	55,529	0.51	0.50	Male
minority	55,529	0.08	0.27	Race: Minority (not Han people)
party	55,529	0.11	0.31	Communist Party member
m_social	55,529	0.08	0.27	Major: Social Sciences
m_natural	55,529	0.11	0.31	Major: Natural Sciences
m_business	55,529	0.14	0.35	Major: Business
m_health	55,529	0.06	0.23	Major: Health
m_engineer	55,529	0.49	0.50	Major: Engineering
m_other	55,529	0.02	0.14	Major: Other major but not Education
onlychild	55,529	0.53	0.50	Only child in the family
urban	55,529	0.50	0.50	Live in urban area
fabaplus	55,529	0.15	0.36	Father's education: BA and above
mobaplus	55,529	0.11	0.31	Mother's education: BA and above
large	55,529	0.30	0.46	Institution size: large
medium	55,529	0.52	0.50	Institution size: medium
inst_985	55,529	0.15	0.36	Institution type: 985 project (extremely selective)
inst_211	55,529	0.17	0.38	Institution type: 211 project (very selective)
inst_univ	55,529	0.47	0.50	Institution type: regular university (average selective)
munici	55,529	0.35	0.48	Institution location: municipalities

Table A3. Propensity Score Matching – Balance Check (US & China)

US					China				
Variable	**Sample**	**STD Diff**	**SD Ratio**	**% Diff**	**Variable**	**Sample**	**STD Diff**	**SD Ratio**	**% Diff**
age	Unmatched	-0.436	0.44		age	Unmatched	0.201	1.44	
	Matched	-0.005	0.62			Matched	-0.007	0.98	
male	Unmatched	-0.251	0.91		male	Unmatched	0.078	1.00	
	Matched	-0.042	0.98	-0.018		Matched	0.008	1.00	0.004
white	Unmatched	0.206	0.92		minority	Unmatched	-0.026	0.96	
	Matched	-0.021	1.01	-0.009		Matched	0.019	1.03	0.005
black	Unmatched	-0.147	0.79		party	Unmatched	0.112	1.14	
	Matched	0.049	1.13	0.010		Matched	-0.022	0.98	-0.008
hispanic	Unmatched	-0.145	0.78		m_social	Unmatched	0.045	1.07	
	Matched	-0.006	0.99	-0.001		Matched	0.019	1.03	0.006
asian	Unmatched	-0.014	0.97		m_natural	Unmatched	-0.03	0.96	
	Matched	-0.018	0.97	-0.004		Matched	0.007	1.01	0.003
m_social	Unmatched	0.121	1.05		m_business	Unmatched	0.060	1.06	
	Matched	-0.022	0.99	-0.010		Matched	-0.010	0.99	-0.004
m_natural	Unmatched	-0.042	0.96		m_health	Unmatched	-0.257	0.62	
	Matched	-0.004	1.00	-0.001		Matched	0.018	1.07	0.003
m_business	Unmatched	-0.027	0.97		m_engineer	Unmatched	-0.028	1.00	
	Matched	-0.004	1.00	-0.001		Matched	-0.012	1.00	-0.006
m_health	Unmatched	-0.214	0.80		m_other	Unmatched	-0.063	0.81	
	Matched	-0.042	0.94	-0.011		Matched	-0.018	0.93	-0.002
m_engineer	Unmatched	-0.060	0.91		onlychild	Unmatched	0.307	0.95	
	Matched	-0.005	0.99	-0.001		Matched	-0.006	1.00	-0.002
m_other	Unmatched	-0.168	0.75		urban	Unmatched	0.238	0.98	
	Matched	-0.028	0.94	-0.005		Matched	-0.009	1.00	-0.005
ft	Unmatched	0.062	0.92		fabaplus	Unmatched	0.251	1.24	
	Matched	-0.018	1.03	-0.005		Matched	0.024	1.01	0.011
trans	Unmatched	-0.376	0.79		mobaplus	Unmatched	0.269	1.36	
	Matched	0.000	1.00	0.000		Matched	0.027	1.02	0.011
dorm	Unmatched	-0.212	0.94		large	Unmatched	-0.079	0.96	
	Matched	-0.021	0.99	-0.010		Matched	-0.005	1.00	-0.002
first	Unmatched	-0.318	0.90		medium	Unmatched	0.086	0.99	
	Matched	0.003	1.00	0.001		Matched	0.012	1.00	0.006
pedu_cl	Unmatched	-0.103	1.00		inst_985	Unmatched	0.023	1.02	
	Matched	-0.011	1.00	-0.005		Matched	0.022	1.02	0.009
pedu_gs	Unmatched	0.274	1.08		inst_211	Unmatched	-0.009	0.99	
	Matched	0.013	1.00	0.007		Matched	-0.014	0.99	-0.006
pri	Unmatched	0.196	1.02		inst_univ	Unmatched	-0.091	0.99	
	Matched	-0.003	1	-0.001		Matched	0.003	1	0.002
medium	Unmatched	0.045	1.03		munici	Unmatched	0.124	1.03	
	Matched	0.053	1.04	0.023		Matched	0.007	1.00	0.004
large	Unmatched	-0.060	0.99						
	Matched	-0.048	0.99	-0.024					
r_neweng	Unmatched	-0.050	0.91						
	Matched	0.032	1.08	0.006					