October 2020

# Estimating the Number of Chinese STEM Students in the United States

CSET Issue Brief



AUTHORS Jacob Feldgoise Remco Zwetsloot

# **Executive Summary**

This paper assesses how many Chinese students are enrolled in U.S. universities across field, degree level, and time. We combine and analyze data from the National Science Board's Science and Engineering Indicators (SEI), the National Center for Science and Engineering Statistics (NCSES), the National Center for Education Statistics (NCES), and the Institute of International Education (IIE).

The results speak to ongoing policy conversations about the risks and benefits of Chinese students enrolled in science, technology, engineering, and mathematics (STEM) programs at U.S. universities. These conversations have been hampered by a lack of granular data on the number of enrolled Chinese students by field and degree level. For example, it is currently impossible to calculate the financial impact of Chinese students on the U.S. university system because we do not know how many Chinese graduate students are in master's programs (and thus likely to pay full tuition) versus Ph.D. programs (for which they often receive university or federal funding). This paper seeks to provide data that can help answer such questions, and to identify remaining data gaps that should be filled.

Because there is no single database of domestic and international students in the United States that includes all the relevant information, analysts have had to produce estimates using several different data sources. These sources often count slightly different things over possibly different periods, complicating the analysis and increasing the risk of inadvertent errors. Our findings differ from widely-cited government estimates. Whereas those estimates suggested that 25 percent of U.S. STEM graduate students and 15 percent of STEM undergraduates are Chinese, we conclude with high confidence that the numbers are 16 percent and 2 percent, respectively.

The key findings are presented in Tables 1 and 2, whose sources and methods are detailed in the Appendix. Table 1 shows the number of enrolled Chinese undergraduate and graduate students by field in academic year 2018/19, focusing on the six main STEM fields at U.S. universities. Across those STEM fields, there are around 46,000 Chinese undergraduates, an estimated 40,000 master's students, and an estimated 36,000 Ph.D. students. At the undergraduate level, around one-third of all Chinese students at U.S. universities are in STEM fields, compared to more than half at the graduate level.

	Number of Chinese Undergraduate	Number of Chinese Graduate Students			
	Students	Overall	Master's ( <i>est.</i> )	Ph.D. ( <i>est.</i> )	
Agricultural Sciences	1,390	1,400	511	889	
Biological Sciences	4,120	5,980	1,585	4,395	
Computer Sciences	11,710	16,990	13,600	3,390	
Engineering	12,890	31,450	16,447	15,003	
Mathematics and Statistics	12,090	12,740	7,436	5,304	
Physical Sciences	3,520	7,500	894	6,606	
Total STEM	45,720	76,060	40,473	35,587	
Total	143,320	129,440	not estimated	not estimated	

Table 1. Number of Chinese Students Enrolled at U.S. Universities by Degree and Field, 2018/19

*Source*: Science and Engineering Indicators; National Center for Science and Engineering Statistics Survey of Graduate Students and Postdoctorates in Science and Engineering. CSET calculated Master's and Ph.D. student estimates. See Appendix A for details.

Table 2 puts the numbers presented in Table 1 into context by calculating the percentage of international and total STEM students at U.S. universities who are Chinese. At the undergraduate level, we find that Chinese students comprise 31 percent of international students and 2 percent of all students at U.S. universities in STEM fields. At the graduate level, we find that Chinese students comprise 37 percent of international students and 16 percent of all U.S. students in STEM fields.

	Chinese Undergraduate Students at U.S. Universities		Chinese Graduate Students at U.S. Universities		
	% of International	% of Total	% of International	% of Total	
Agricultural Sciences	44%	1%	32%	10%	
Biological Sciences	24%	1%	36%	7%	
Computer Sciences	33%	3%	26%	14%	
Engineering	23%	2%	37%	19%	
Mathematics and Statistics	72%	11%	64%	33%	
Physical Sciences	42%	3%	40%	14%	
Total STEM	31%	2%	37%	16%	

Table 2. Chinese Students as a Percentage of U.S. Undergraduate andGraduate STEM Students, 2018/19

*Source*: Science and Engineering Indicators; National Center for Science and Engineering Statistics Survey of Graduate Students and Postdoctorates in Science and Engineering; National Center for Education Statistics. See Appendix A for details.

The remainder of this paper discusses why these numbers matter and how they were estimated; provides historical trend data; and identifies remaining data gaps on Chinese students and researchers in the United States that agencies and researchers should fill.

# Background

Policymakers, the national security community, and academia are engaged in a heated debate about the risks and benefits of having students from China enrolled in U.S. universities. Some refer to all Chinese students as potential threats,<sup>1</sup> while FBI documents and policymaker statements on the risks posed by Chinese students often place special emphasis on "post-graduate students and post-doctorate researchers studying [STEM]."<sup>2</sup> Policymakers and researchers also ask whether U.S. universities have become financially dependent on Chinese students and how much Chinese students contribute to U.S. innovation.<sup>3</sup>

This paper attempts to distinguish students, where possible, by their degree level and field. Since "STEM" is a broad concept, data is presented by specific academic field. We also attempt to distinguish between master's and Ph.D. students within the category of "graduate students." We make this distinction for three reasons. First, master's students may be less of an innovation asset and therefore less of a security concern than Ph.D. students.<sup>4</sup> Second, master's students often pay full tuition, whereas Ph.D. students tend to be funded by universities or federal research grants, meaning the two groups are different from an economic perspective.<sup>5</sup> Third, there is data on how many Ph.D. students stay in the United States after graduating but not on how many master's students stay; to estimate how many Chinese students stay in total, we need to know the relative sizes of these two groups.<sup>6</sup>

However, because all the relevant datasets have informational gaps, estimating Chinese enrollments across fields and degree level is not straightforward (Box 1).

## Box 1. Data Sources on International Student Enrollments in the United States

This paper draws on four datasets to estimate the number of Chinese STEM students at U.S. universities:

- Science and Engineering Indicators' Student and Exchange Visitor Information System data tables. The Department of Homeland Security maintains SEVIS to track students and exchange visitors at U.S. educational institutions. The Science and Engineering Indicators team has published SEVIS-derived data tables on international undergraduate and graduate enrollments by citizenship and field since 2005/06.<sup>7</sup>
- National Center for Science and Engineering Statistics' Survey of Graduate Students and Postdoctorates in Science and Engineering. The NCSES GSS surveys departments about graduate enrollments in science, engineering, and health fields, broken out by domestic and international students. Its data goes back decades, but it has only distinguished between master's and Ph.D. students since 2017/18.<sup>8</sup>
- National Center for Education Statistics Integrated Postsecondary Education Data System. IPEDS tracks enrollments and degree completions at U.S. universities by degree level, citizenship (domestic versus international), and field.<sup>9</sup> Its data coverage varies, but most information is available back to the 1990s.
- Institute of International Education's Open Doors survey. In its annual Open Doors survey, IIE asks U.S. universities to report the number of enrolled international students by degree level, by field, and by country of origin.<sup>10</sup> Data on Chinese students is available at least back to 1999/2000.

These datasets differ along several dimensions. Some distinguish by specific nationality (SEI SEVIS, IIE Open Doors) whereas others only provide data on international students broadly (NCSES GSS, NCES IPEDS). Some distinguish between master's and Ph.D. students (e.g. NCSES GSS) whereas others lump them together under the "graduate student" heading (e.g. SEI SEVIS). Some allow for simultaneous disaggregation by citizenship, degree level, and field (e.g. SEI SEVIS), whereas others only break down across one or two dimensions at a time (e.g. IIE Open Doors).

Box 1, continued.

Different datasets may also cover different time periods and use different field categorizations (see Appendix B).

These strengths and weaknesses mean that information from different datasets must be combined in order to derive specific estimates of Chinese (and other international) enrollments across degree level and field. Appendix A discusses our approach in more detail.

# Comparison with Prior Estimates

Because centralized data sources are lacking, analysts have had to estimate the proportion of students at U.S. universities who are Chinese. Perhaps the most widely-cited estimate is found in a prominent 2018 report by the Defense Innovation Unit Experimental (DIUx)<sup>\*</sup> titled *China's Technology Transfer Strategy: How Chinese Investments in Emerging Technology Enable A Strategic Competitor to Access the Crown Jewels of U.S. Innovation.*<sup>11</sup> The DIUx report "infer[red] that 25% of [U.S.] graduate students in STEM fields are Chinese foreign nationals," an estimate that several other government documents have since cited.<sup>12</sup>

The DIUx estimate differs from our calculations, which show that 16 percent of graduate students in STEM fields are Chinese foreign nationals (Table 2). In this section, we briefly explain the difference between our calculations and the DIUx report's estimate. This discussion is meant to clarify data issues, not to impugn the report or dispute its broader conclusions. Indeed, as discussed in the Conclusion, we believe one takeaway from this study is that federal agencies should collect and publish data that is sufficiently detailed and disaggregated to eliminate the need for analyses such as these in the first place.

<sup>&</sup>lt;sup>\*</sup> In 2018, DIUx dropped "experimental" from its name, and is now known as DIU. Because the report was published when it was DIUx, we refer to the "DIUx report" throughout this paper.

The DIUx report's estimate of 25 percent appears based on its claim that "For doctoral programs, 57% of engineering, 53% of computer science and 50% of math and statistics candidates were foreign; half of these were Chinese."<sup>13</sup> There are two problems with basing an estimate based on this claim. First, master's students and several STEM fields are excluded. Second, except in mathematics, NSF data shows that fewer than half of international U.S. Ph.D. graduates in these fields are Chinese.<sup>14</sup> The DIUx estimate appears to be based on a misreading of a news article.<sup>\*</sup>

The DIUx report also claims that "45% of STEM undergraduates are foreign and [one-third] of them are from China." This implies that 15 percent of U.S. STEM undergraduates are Chinese. However, according to IPEDS data, only 7 percent of undergraduate STEM degrees awarded in 2018/19 went to international students, and, as shown in Table 2, we found that only 2 percent of U.S. STEM undergraduates (about one-third of 7 percent) are Chinese citizens.<sup>†</sup>

<sup>&</sup>lt;sup>\*</sup> The source DIU cites is a 2015 article by *Pew Research*. The *Pew* article does not claim that half of international Ph.D. graduates are Chinese. The closest statement we were able to find in the article says that "China ... is the single biggest source of foreign college students, and most of the increase over the past decade and a *half can be attributed to Chinese students*." The italicized portion of the sentence (our italics) may have been misinterpreted as saying that half of international Ph.D. students are Chinese. See Drew Desilver, "Growth from Asia Drives Surge in U.S. Foreign Students," *Pew Research Center*, June 18, 2015, https://www.pewresearch.org/fact-tank/2015/06/18/growth-from-asia-drives-surge-in-u-s-foreign-students/.

<sup>&</sup>lt;sup>†</sup> This DIU estimate also appears to be based on a misinterpretation of a secondary source. In its sentence on undergraduates, DIUx cited a 2014 article in *Science*. The 45 percent claim from DIUx appears to be based on a line in the *Science* article that says, referring to international students, "Those pursuing undergraduate degrees in so-called STEM fields make up 45% of the undergraduate total." However, the *Science* statistic refers to the proportion of international undergraduates who enroll in STEM fields (as opposed to in non-STEM fields), not the proportion of U.S. STEM undergraduates who are international. DIU's claim that one-third of international undergraduates in the United States are Chinese was correct (see Table 2), and that statistic was also mentioned in the *Science* article. See Jeffrey Mervis, "Data Check: Why Do Chinese and Indian Students Come to U.S. Universities?", *Science*, November 18, 2014, https://www.sciencemag.org/news/2014/11/datacheck-why-do-chinese-and-indian-students-come-us-universities.

# Trends: Chinese Enrollments Over Time

The first wave of Chinese students that came to the United States in the 1980s, during China's reform period, consisted mainly of graduate students.<sup>15</sup> Chinese enrollments rose rapidly starting in the late 2000s, as China's economy boomed.<sup>16</sup> Figure 1 shows that the number of Chinese students in the United States grew from around 60,000 in 2005/06 to around 370,000 in 2018/19. The figure distinguishes between four categories of students: graduate, undergraduate, Optional Practical Training participants, and "other."<sup>\*</sup> Enrollment growth in the past decade is mainly driven by growth in undergraduate enrollments and OPT participants, although graduate enrollments also roughly doubled during this period. In 2018/19, there were approximately 148,900 Chinese undergraduate students, 133,400 graduate students, 70,000 OPT participants, and 17,200 other students at U.S. universities.

<sup>&</sup>lt;sup>\*</sup> The Optional Practical Training (OPT) program allows F-1 student visa holders to work in the United States for up to one year (for non-STEM students) or three years (for STEM students). The "other" category consists primarily of students in non-degree programs (e.g. intensive English language and certificate programs) and exchange students. Before 2006/07, IIE included OPT students in the "other" category.



**Figure 1.** Chinese enrollments at U.S. universities by degree level and program, 2000/01 to 2018/19.

Figures 2 and 3 show STEM fields of study for the undergraduate and graduate student populations, respectively, using SEI SEVIS data (available from 2005/06 onward). Figure 2 shows that enrollment in every STEM field grew rapidly between 2006/07 and 2016/17, especially after 2008/09. Chinese undergraduate enrollment grew most rapidly in Engineering, Mathematics and Statistics, and Computer Sciences.



**Figure 2**. Number of Chinese undergraduate students at U.S. universities in STEM fields, 2005/06 to 2018/19.

*Source*: Science and Engineering Indicators. Data for 2015/16 is missing and was imputed as the 2014/15-2016/17 average.

Figure 3 shows that from 2005/06 to 2018/19, Chinese graduate enrollment grew most rapidly in Engineering, Mathematics and Statistics, and Computer Sciences, as was the case with undergraduate enrollment. In contrast, Chinese graduate enrollment in Biological and Physical Sciences remained nearly constant over the same span.<sup>\*</sup>

<sup>&</sup>lt;sup>\*</sup> It is not possible to analyze historical trends for Chinese master's and Ph.D. students specifically because field-specific master-to-Ph.D. enrollment ratios are not available from the NCSES GSS until 2017/18. This data therefore groups all graduate students together. See Appendix B for a discussion.



**Figure 3**. Number of Chinese graduate students at U.S. universities in STEM fields, 2005/06 to 2018/19.

*Source*: Science and Engineering Indicators. Data for 2015/16 is missing and was imputed as the 2014/15-2016/17 average.

## Chinese Shares of U.S. STEM Enrollments

Table 2 showed what shares of international and total STEM students at U.S. universities were Chinese students in 2018/19. This is the most recent snapshot we have. The following section uses data from SEI's SEVIS tables to assess how these shares changed over time. It first looks at the undergraduate level, and then at the graduate level.

#### Undergraduate Enrollments

Figure 4 shows that the Chinese share of international undergraduates has grown steadily. Across all STEM fields, Chinese students went from accounting for 4 percent of international undergraduate enrollments in 2005/06 to accounting for 31 percent in 2018/19. There is notable variation across fields; in Mathematics and Statistics, their share grew from 9 percent to 72 percent, while in Biological Sciences their share grew from 3 percent to 24 percent. These rising shares indicate that in this period—and especially in the few years before and after 2010/11—Chinese undergraduate enrollments increased much faster than non-Chinese international enrollments.



Figure 4. Chinese undergraduate enrollments as a percentage of

*Source*: Science and Engineering Indicators. Data for 2015/16 is missing and was imputed as the 2014/15-2016/17 average. See Appendix A for details.

Figure 5 shows that across all STEM fields, Chinese undergraduate students as a percentage of total U.S. undergraduate students—domestic as well as international—grew from 0.1 percent in 2005/06 to 2 percent in 2018/19. However, percentages in some fields grew more rapidly than others. Chinese students in Mathematics and Statistics as a percent of total undergraduate students in this field grew from 0.3 percent to 11 percent, while Chinese students in Biological Sciences only grew from 0.1 percent to 0.6 percent.



**Figure 5**. Chinese undergraduate enrollments as a percentage of total undergraduate enrollments by STEM field, 2005/06 to 2018/19.

*Source*: Science and Engineering Indicators; National Center for Education Statistics. Data for 2015/16 is missing and was imputed as the 2014/15-2016/17 average. See Appendix A for details.

#### Graduate Enrollments

Figures 6 and 7 replicate these same calculations at the graduate level. Figure 6 shows that Chinese enrollment shares grew less quickly at the graduate level than at the undergraduate level, mainly because graduate enrollment shares were already higher in 2005/06–27 percent in all of STEM, compared to 0.1 percent at the undergraduate level. Between 2012/13 and 2018/19, Chinese graduate enrollment shares among international students remained around 35 percent. This shows that during this period non-Chinese international enrollments were keeping pace with Chinese enrollments, even as the number of Chinese graduate students increased from around 50,000 in 2012/13 to more than 70,000 in 2018/19 (as seen in Figure 3).





*Source*: Science and Engineering Indicators. Data for 2015/16 is missing and was imputed as the 2014/15-2016/17 average. See Appendix A for details.

Figure 7 shows that similar trends characterize Chinese graduate student enrollment as a percent of total U.S. graduate enrollment. Across all STEM fields, Chinese students as a percentage of total graduate students grew from 9 percent in 2005/06 to 16 percent in 2018/19, with most of the increase happening prior to 2012/13. The main exception is Mathematics and Statistics, where Chinese enrollment growth consistently outpaced domestic and other international enrollment growth between 2007/08 and 2017/18.



**Figure 7**. Chinese graduate enrollments as a percentage of total graduate enrollments by STEM field, 2005/06 to 2018/19.

*Source*: Science and Engineering Indicators; National Center for Science and Engineering Statistics. Data for 2015/16 is missing and was imputed as the 2014/15-2016/17 average. See Appendix A for details.

# Conclusion

This paper used four different data sources to analyze how many Chinese STEM students are enrolled at U.S. universities. We found that there were an estimated 46,000 Chinese undergraduates, 40,000 master's students, and 36,000 Ph.D. students from China in U.S. STEM programs in 2018/19. They accounted for 2 percent of total U.S. STEM enrollments at the undergraduate level and 16 percent at the graduate level.

These numbers are highly relevant to ongoing debates about the risks and benefits to the United States of hosting Chinese students and researchers. These risks and benefits differ across fields and degree levels—master's students, for example, are distinct from Ph.D. students in both the work they do and whether they are a source of revenue for universities. Data alone will not solve these debates, which involve difficult trade-offs, but it is necessary for formulating risk-management strategies that do not unnecessarily harm U.S. universities and innovation. For example, the data presented in this report can help policymakers assess how many Chinese students would be affected by several visa screening bills currently under consideration in Congress. We had to draw on several datasets in putting together this report because no single source contained all the relevant information, and in distinguishing between master's and Ph.D. students, we had to make estimates because no data source was sufficiently granular. The U.S. government should address these data issues in order to obviate the need for estimates and eliminate the risk of miscalculations. Appendix B highlights specific gaps across all the relevant datasets in order to help statistical agencies collect and publish sufficiently detailed data in the future.

# Acknowledgments

For feedback and assistance, we would like to thank Julie Baer, Dale Brauner, Michael Brown, James Dunham, Alex Friedland, Daniel Hague, Igor Mikolic-Torreira, Dahlia Peterson, Pavneet Singh, Josh Trapani, Lynne Weil, and Emily Weinstein.



© 2020 by the Center for Security and Emerging Technology. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

To view a copy of this license, visit https://creativecommons.org/licenses/by-nc/4.0/.

Document Identifier: doi: 10.51593/20200023

Center for Security and Emerging Technology | 18

# Appendix A: Methodology

# Sources and Methods for Table 1

Table 1 consists of a combination of administrative data and estimates. For undergraduate and graduate students as a whole, we can use administrative data from the DHS SEVIS system published through the SEI.<sup>17</sup> This data is displayed in the first two columns of Table 1. The third and fourth columns of Table 1 list the numbers of Chinese master's and Ph.D. students across fields. Because there is no public data on the number of Chinese students across graduate degree levels, we had to estimate these numbers. This section lays out the assumptions and sources that we used to produce our estimates.

We derive our estimates by (1) calculating the ratio of master's to Ph.D. students per field for all international students, and (2) assuming that these ratios are the same for Chinese students as they are for all international students. The data source that can be used for these calculations is the NCSES's GSS survey, which tracks the number of enrolled international master's and Ph.D. students by field within science and engineering disciplines.<sup>18</sup> According to NCSES GSS data, there are, for example, roughly four international master's students for every international Ph.D. student in computer science. Similar ratios can be calculated for other fields.<sup>\*</sup> Next, we apply these ratios to the Chinese graduate student totals found in the SEVIS data tables. For example, out of the 16,990 Chinese graduate students in computer science in 2018/19, we estimate that around four out of five (13,600) are master's students and one out of five (3,390) are Ph.D.

These estimates are sure to have some error. How much error they have depends on how different Chinese graduate students are from the average international graduate student in their choice to enroll in master's versus Ph.D. programs. The more similar Chinese graduate students are to the average

<sup>\*</sup> According to the GSS, the percentage of international graduate students who are master's students is about 52.3 percent in engineering, 36.5 percent in the agricultural sciences, 26.5 percent in the biological sciences, 80.1 percent in computer science, 58.4 percent in mathematics and statistics, and 11.9 percent in the physical sciences.

international graduate student, the less error there will be. We were able to locate only one study on this question, which shows that Chinese students' distribution across master's and Ph.D. degrees is within a few percentage points of the distribution among all international students.<sup>19</sup> However, this data is not STEM-specific. More data would be needed to fully assess the magnitude and direction of any error in our method.

This method cannot be applied to non-STEM fields because the NCSES GSS only tracks enrollments in science and engineering disciplines. The bottom cells in the third and fourth column of Table 1 therefore contain no estimates.

#### Sources and Methods for Table 2, Figures 4-7

Table 2 estimates the proportion of undergraduate and graduate students at U.S. universities who are Chinese nationals, and Figures 4-7 showed trends for these proportions over time. This required comparing the number of Chinese students to the number of international students and the overall number of students. Because data on U.S. enrollments is not tracked in a centralized way, we used three different sources to calculate the percentages.

Data on Chinese students comes from SEVIS tables published by the SEI, the same source used for Table 1 (discussed above). These same SEVIS tables also contain administrative data on the total number of international students, so the percentage of *international* students who are Chinese in each of our STEM fields can be calculated using the SEVIS tables.

However, SEVIS only tracks international students, not domestic students, so different sources need to be used to calculate the percentage of *total* students who are Chinese (the second and fourth columns). At the graduate level, NCSES GSS data can be used to calculate the proportion of all graduate students who are international. This proportion (international of total) can then be combined with the proportion calculated from the SEVIS data (Chinese of international) to estimate what share of all graduate students are Chinese. For example, SEVIS data tells us that 26 percent of international graduate students in computer science are Chinese, and the GSS tells us that 55 percent of all U.S. computer science graduate students are international. Multiplying these two percentages together, we estimate that about 14 percent of all U.S. computer science graduate students are Chinese. Table 3 illustrates this process.

	(1) International Students as % of U.S. Total		(2) Chinese Students as % of International		(3) Chinese Students as % of U.S. Total
Agricultural Sciences	30%		32%		10%
Biological Sciences	20%		36%		7%
Computer Sciences	55%	x	26%	=	14%
Engineering	53%		37%		19%
Mathematics and Statistics	51%		64%		33%
Physical Sciences	35%		40%		14%
Total STEM	44%		37%		16%

**Table 3.** Methodology for calculating the share of all U.S. STEM studentswho are Chinese at the graduate level.

*Sources*: Science and Engineering Indicators; National Center for Science and Engineering Statistics. Percentages were calculated for 2018/19.

We use broadly the same procedure at the undergraduate level. Because the NSF only tracks graduate students, for undergraduate numbers we use the National Center for Education Statistics' Integrated Postsecondary Education System (IPEDS). IPEDS tracks the number of *enrolled* undergraduate students in Engineering, Physical Sciences, Mathematics and Statistics, and Biological Sciences; for the Computer Sciences and Agricultural Sciences, we have to rely on IPEDS data on *degrees granted*.<sup>20</sup> With this data, we can again multiply the share of international undergraduate students who are Chinese (obtained from SEVIS data) with the share of all bachelor's degrees that are granted to international students (obtained from IPEDS) to estimate the proportion of all students who are Chinese.

	(1) International Students as % of U.S. Total		(2) Chinese Students as % of International		(3) Chinese Students as % of U.S. Total
Agricultural Sciences	3%		44%		1%
Biological Sciences	3%		24%		1%
Computer Sciences	9%	x	33%	=	3%
Engineering	9%		23%		2%
Mathematics and Statistics	16%		72%		11%
Physical Sciences	6%		42%		3%
Total STEM	7%		31%		2%

**Table 4.** Methodology for calculating the share of all U.S. STEM studentswho are Chinese at the undergraduate level.

Sources: Science and Engineering Indicators; National Center for Education Statistics. Percentages were calculated for 2018/19.

These steps allowed us to estimate field-specific proportions. To produce an aggregate STEM estimate, these field-specific proportions need to be combined. Some STEM fields are much larger than others in terms of enrollment, and these differing sizes should be taken into account. For example, 37 percent of graduate degrees granted in STEM fields in 2018/19 were in Engineering disciplines, whereas only 7 percent were in Mathematics and Statistics. The proportions for all U.S. STEM students are thus calculated as weighted averages, where the weights are based on overall national enrollment numbers. These weighted percentages are listed in the bottom rows of Tables 3 and 4.

# Appendix B: Data Gaps and Future Work

There are several remaining questions about Chinese STEM enrollments that this Brief was unable to answer.

- We were unable to show detailed field breakdowns (e.g. "Chemistry" and "Physics" instead of "Physical Sciences") because SEI's SEVIS data tables and most other datasets do not break out by fine field.
- We were unable to estimate the number of Chinese master's and Ph.D. students in most non-STEM sub-fields because the NCSES GSS and SEI's SEVIS data tables only track enrollments in science and engineering fields.
- We were unable to estimate what proportion of U.S. master's and Ph.D. students specifically (as opposed to graduate students broadly) are accounted for by Chinese students because SEI's SEVIS tables only refer to "graduate students."
- We were unable to estimate historical trends for Chinese master's and Ph.D. enrollments because the NCSES GSS has only distinguished between master's and Ph.D. students since 2017/18.

We point out these specific data gaps in the hope that future data collection and aggregation efforts by federal agencies can fill them. The main existing source that could fill these gaps is the DHS SEVIS system, which contains detailed individual-level data on all international students. SEVIS data is currently not accessible to researchers, but National Science Board staff do appear to have access and may be able to publish additional data tables that contain some of this information.<sup>21</sup> DHS SEVIS data could also be used to assess the accuracy of the master's and Ph.D. estimates presented in Table 1.

Future work could also extend this analysis to the post-doctoral level and to other countries. Due to data gaps, it is even more difficult to estimate the number of Chinese postdocs and visiting researchers doing STEM research in the United States; this question will require a separate analysis. The data sources and methods used in this paper (detailed in Appendix A) could also be extended to other common countries of origin among international students, such as India or South Korea.

# Endnotes

<sup>1</sup> Bill Gertz, "China Using Students as Spies," *Washington Times*, April 25, 2018, <u>https://www.washingtontimes.com/news/2018/apr/25/china-uses-students-as-spies/</u>.

<sup>2</sup> Federal Bureau of Investigations, "China: The Risk to Academia," 2019, 2, <u>https://www.fbi.gov/file-repository/china-risk-to-academia-2019.pdf/view</u>; Gregg Re, "Tom Cotton Suggests Chinese Students Shouldn't Be Allowed to Study Sciences in the US," *Fox News*, April 26, 2020, <u>https://www.foxnews.com/politics/tom-cotton-suggestschinese-students-shouldnt-be-allowed-to-study-sciences-in-the-us.</u>

<sup>3</sup> Makala Skinner, "The Financial Risk of Overreliance on Chinese Student Enrollment," *World Education News + Reviews*, December 17, 2019, <u>https://wenr.wes.org/2019/12/the-financial-risk-of-overreliance-on-chinese-student-</u> <u>enrollment</u>; Patrick Gaule and Mario Piacentini, "Chinese Graduate Students and U.S. Scientific Productivity," *The Review of Economics and Statistics* 95, no. 2 (May 2013): 698-701, <u>https://www.mitpressjournals.org/doi/abs/10.1162/REST\_a\_00283</u>

<sup>4</sup> For example, doctoral graduates are more likely to work in R&D jobs than master's graduates; see National Science and Engineering Indicators, "Scientists and Engineers Performing Research and Development, Table 3-7," September 2019, https://ncses.nsf.gov/pubs/nsb20198/s-e-workers-in-the-economy.

<sup>5</sup> Exact statistics on how many Chinese master's and Ph.D. students are self-funded is not available, but experts and available data strongly suggest that master's students are much more likely to self-fund than Ph.D. students. According to the NSF Survey of Graduate Students and Postdoctorates in Science and Engineering, 91 percent of STEM Ph.D. students are supported by either institutional (university) funds or by federal research funds in 2018/19, while 70 percent of master's students in STEM fields were self-funded. This data source does not distinguish between domestic and international students, however; see National Center for Science and Engineering Statistics, "Survey of Graduate Students and Postdoctorates in Science and Engineering, Fall 2018," Table 3-1, https://ncsesdata.nsf.gov/gradpostdoc/2018/html/gss18-dt-tab003-1.html. The Institute for International Education, which looks specifically at international students, reports funding sources for "graduate" students as a whole. For that group, 58 percent are self-funded, 36 percent are funded by a U.S. university, and 3 percent are funded by a foreign government or university; see Institute for International Education, "Open Doors: Primary Source of Funding," https://opendoorsdata.org/infographic/primary-source-of-funding-ofinternational-students/

<sup>6</sup> For a discussion of data sources on stay rates, see Remco Zwetsloot, James Dunham, Zachary Arnold, and Tina Huang, "Keeping Top AI Talent in the United States" (Center for Security and Emerging Technology, December 2019), <u>https://cset.georgetown.edu/wp-content/uploads/Keeping-Top-AI-Talent-in-the-United-States.pdf</u>.

<sup>7</sup> National Science Board, "Higher Education in Science and Engineering: Data," <u>https://ncses.nsf.gov/indicators/data#data-sources-tab</u>. Data for 2015 is missing.

<sup>8</sup> National Center for Science and Education Statistics, "Survey of Graduate Students and Postdoctorates in Science and Engineering," <u>https://www.nsf.gov/statistics/srvygradpostdoc/</u>.

<sup>9</sup> National Center for Education Statistics, "Integrated Postsecondary Education Data System: About IPEDS," <u>https://nces.ed.gov/ipeds/about-ipeds.</u>

<sup>10</sup> Institute for International Education, "Open Doors: FAQ," <u>https://opendoorsdata.org/faq/</u>.

<sup>11</sup> Michael Brown and Pavneet Singh, "China's Technology Transfer Strategy: How Chinese Investments in Emerging Technology Enable A Strategic Competitor to Access the Crown Jewels of U.S. Innovation" (Defense Innovation Unit Experimental, January 2018), 18, <u>https://admin.govexec.com/media/diux\_chinatechnologytransferstudy\_jan\_2018\_(1).pdf</u>.

<sup>12</sup> See, e.g., Office of Trade and Manufacturing Policy, *How China's Economic Aggression Threatens the Technologies and Intellectual Property of the United States and the World* (Washington, DC: White House, June 2018), 14, <u>https://www.whitehouse.gov/wp-content/uploads/2018/06/FINAL-China-Technology-Report-6.18.18-PDF.pdf</u>; Office of the Deputy Assistant Secretary of Defense for Industrial Policy, *Assessing and Strengthening the Manufacturing and Defense Industrial Base and Supply Chain Resiliency of the United States* (Washington, DC: Department of Defense, September 2018), 44, <u>https://media.defense.gov/2018/Oct/05/2002048904/-1/-1/1/ASSESSING-AND-STRENGTHENING-THE-MANUFACTURING-AND-DEFENSE-INDUSTRIAL-BASE-AND-SUPPLY-CHAIN-RESILIENCY.PDF</u>.

<sup>13</sup> The DIUx report also cites other data prior to its claim that 25 percent of U.S. graduate students are Chinese, including statistics on patenting by universities, but these statistics do not seem germane to our question. The DIUx report never explicitly spells out its methodology for the estimate, so this section represents our best attempt to reverse engineer its process.

<sup>14</sup> Remco Zwetsloot, Jacob Feldgoise, and James Dunham, "Trends in U.S. Intention-to-Stay Rates of International Ph.D. Graduates Across Nationality and STEM Fields" (Center for Security and Emerging Technology, April 2020), <u>https://cset.georgetown.edu/wpcontent/uploads/CSET-Trends-in-U.S.-Intention-to-Stay-Rates.pdf</u>. Table 1 in that paper displays the number of U.S. Ph.D. graduates by country of origin and field (averaged across 2015-2017). These numbers indicate that in engineering, 37 percent of international Ph.D. graduates are Chinese (22 percent of the U.S. total); in computer science, 37 percent of international Ph.D. graduates are Chinese (24 percent of the U.S. total); and in mathematics, 49 percent of international Ph.D. graduates are Chinese (25 percent of U.S. total). These are the three fields with the highest concentration of Chinese and international Ph.D. students; other STEM fields in the life sciences and physical sciences are less internationalized. Averaged across all STEM fields, Chinese Ph.D. graduates constitute 16 percent of U.S. STEM Ph.D. graduates. That statistic was calculated using data from the NSF Survey of Earned Doctorates. The NSF SED data supports the findings in Table 2 of this report, which is based on different data sources and which concern all Chinese graduate students instead of only Ph.D. graduates, but which also arrives at 16 percent as the STEM-wide average enrollment share.

<sup>15</sup> Small numbers of Chinese students came to the United States prior to the 1970s as well, but numbers increased drastically in the 1980s. For an account of the early years of Chinese students in the United States, see, e.g., Leo Orleans, *Chinese Students in America: Policies, Issues, and Numbers* (Washington, DC: National Academies Press, 1988), https://www.nap.edu/catalog/743/chinese-students-in-america-policies-issues-andnumbers.

<sup>16</sup> For background on Chinese students in the United States, see Yingyi Ma, *Ambitious and Anxious: How Chinese College Students Succeed and Struggle in American Higher Education* (New York: Columbia University Press, 2019). Explanations for the enrollment boom since the mid-2000s include the emergence of China's middle class, the relative shortage of high-quality universities in China, budget shortfalls at U.S. universities since the 2008 financial crisis, and the appreciation of the renminbi (China's currency) relative to the dollar.

<sup>17</sup> Science and Engineering Indicators, "Higher Education in Science and Engineering: Tables," September 2019, <u>https://ncses.nsf.gov/pubs/nsb20197/data</u>, Tables S2-13 (for undergraduate students) and S2-14 (for graduate students).

<sup>18</sup> National Center for Science and Education Statistics, "Survey of Graduate Students and Postdoctorates in Science and Engineering: Table 2-4," <u>https://ncsesdata.nsf.gov/gradpostdoc/2018/html/gss18-dt-tab002-4.html</u>. The GSS does not track the specific nationalities of international students. It has distinguished between master's and Ph.D. students only since 2017/18 (before that it tracked the number of "graduate students"). The GSS defines "science and engineering" to include STEM fields but also the social sciences.

<sup>19</sup> The Council on Graduate Schools (CGS) tracks international student enrollments by country of citizenship in its annual International Graduate Admissions Survey, which distinguishes between master's and Ph.D. students. In its most recent survey, CGS found that 36 percent of Chinese graduate students enrolled in a Ph.D., compared to 41 percent among

all international graduate students who enrolled in Fall 2019. See Enyu Zhou, Radomir Ray Mitic, Christian P.L. West, and Hironao Okahana, "International Graduate Applications and Enrollment: Fall 2019" (Council of Graduate Schools, February 2020), Table B4, <a href="https://cgsnet.org/sites/default/files/civicrm/persist/contribute/files/CGS%20Fall%202">https://cgsnet.org/sites/default/files/civicrm/persist/contribute/files/CGS%20Fall%202</a> 019%20International%20Report.pdf.

<sup>20</sup> IPEDS appears to track enrollments for Engineering, Physical Sciences, Mathematics and Statistics, and Biological Sciences only for even-numbered years (e.g. 2006, 2008, etc.); for odd years, we used the data point from one year prior. For the Computer Sciences and the Agricultural Sciences, IPEDS appears not to track any enrollments data, so we use data on annual graduates instead. For any given year, the proportion of bachelor's degree *graduates* who are international students may not be exactly the same as the proportion of *enrolled* students who are international. The class that enrolled in 2015, for example, should be most demographically similar to the graduating class of 2019. To mitigate this source of error, for each enrolled year where we lacked enrollment data, we used the data on degrees granted from their anticipated graduation year (four years later). For classes enrolled in 2015 and later, we used the most recent IPEDS data from 2019. These issues with IPEDS may introduce some error into our estimates, but the proportion of international students at the undergraduate level is relatively low and has remained fairly stable over time, so any error is likely to be small.

<sup>21</sup> SEVIS does have a data portal where researchers can interact with the data through a user interface. However, the portal does not allow for disaggregation beyond the levels already covered by the datasets used in this paper. See "SEVIS Data Mapping Tool," *Study in the States*, https://studyinthestates.dhs.gov/sevis-data-mapping-tool.