DOE/OR/00033--T650

Foreign Nationals Who Receive Science or Engineering Ph.D.s from U.S. Universities: Stay Rates and Characteristics of Stayers

Prepared by:

Michael G. Finn
Leigh Ann Pennington
Kathryn Hart Anderson

Analysis and Evaluation Programs
Science/Engineering Education Division
Oak Ridge Institute for Science and Education

April 1995

This document describes activities performed under Contract Number DE-AC05-76OR00033 between the U.S. Department of Energy and Oak Ridge Associated Universities. Research for this document was supported by a grant from the Alfred P. Sloan Foundation to Oak Ridge Associated Universities and also by a grant from the National Science Foundation to Oak Ridge Institute for Science and Education.

All opinions expressed in this document are the authors' and do not necessarily reflect policies and views of the U.S. Department of Energy or the Oak Ridge Institute for Science and Education.

TABLE OF CONTENTS

LIST OF TABLES

Summary of Findings

- Less than one-half of the foreign students who earned Ph.D. degrees in science or engineering during the1980s were working in the United States in 1992. The proportion was around 41 to 42 percent of those who were on temporary visas at graduation, and 48 to 49 percent of all, including those who had permanent resident visas at the time of graduation.

- The stay rate varied by degree field, with engineering having the highest stay rate, and social sciences and life sciences the lowest.

- The stay rate varied greatly by country of citizenship. Students from India, the Peoples Republic of China, and Iran showed stay rates that were well above average. Students from Korea, Japan and Brazil showed stay rates well below average.

- Foreign students attaining S&E Ph.D.s in 1990 seem to be staying at a somewhat higher rate than students from the 1980s. This could be explained by the increasing number of students from the Peoples Republic of China and the high stay rate those students have experienced since 1990.

- Foreign students who leave are as likely to have graduated from a highly rated department as are those who stay.

- The work history of foreign students who received Ph.D.s from U.S. universities from 1981 thru 1986 indicated that productivity differences (as measured by salaries) between those who stayed and those who left by 1991 did not influence the decision to emigrate. However, those who worked for nonprofit firms or government organizations in the U.S. were more likely to leave than their cohorts working in industry or in university jobs.

Overview

This report studies the behavior of foreign nationals who received Ph.D. degrees in science or engineering from U.S. universities during the period 1984-1990. It addresses two distinct questions:

What proportion of foreign students stay to work in the United States after graduation?

Do foreign students who leave the United States differ from those who stay?

Chapter 1 provides descriptive statistics to answer the first question. These estimates of stay rates have small margins of error because they were produced from the tax payment records of the Social Security Administration. The estimates of stay rates in Chapter 1 also provide a partial answer to the second question as well as we are able to provide stay rates for different degree fields and different countries of citizenship, thereby identifying country-specific and field-specific differences in stay rates.

More information on the differences between leavers and stayers is provided in Chapter 2. In this section, we conduct a statistical analysis of respondents to the Survey of Doctorate Recipients. While this forces us to deal with a relatively small sample of foreign students, we are able to focus on those who work in the United States for a few years and then leave. We compare these foreign graduates with others who stay here for a longer period of time and identify factors more frequently associated with leavers.

Chapter 1
FOREIGN NATIONAL STAY RATES

This chapter provides estimates of stay rates for a large number of groups of foreign nationals who received doctorates in science and engineering (S&E) from U.S. universities. Each line in the several tables that follow describes a different group of these degree recipients. We will describe the nature of our estimates and some of their qualifications in a discussion of Table 1-1. For example, the first line of Table 1-1 provides stay rates for 987 persons who were foreign nationals here on temporary resident visas at the time they earned Ph.D.s in the physical sciences in 1984. We made these estimates by drawing a sample of 200 of the 987 persons from the Doctorate Records File and requesting the Social Security Administration to calculate the proportion of these who recorded at least $5,000 in earnings in social security covered

Table 1-1. Temporary Residents Receiving S&E Ph.D.s in 1984 Who Were Working in the United States, 1986-1992

Degree Field[1]	Temp. Res. Ph.D.s	Percent Working in United States						
		1986	1987	1988	1989	1990	1991	1992
Physical Sciences	987	46%	48%	48%	48%	47%	46%	45%
Life Sciences	673	24	27	31	31	32	34	33
Social Sciences	746	26	28	25	25	27	28	26
Engineering	1,247	52	54	55	56	56	56	55
Aeronautical Eng.	57	65	67	71	71	65	62	64
Ag. & Bio. Eng.	55	26	20	29	29	29	25	23
Chemical Eng.	145	59	62	62	65	64	67	65
Civil Eng.	162	47	45	45	45	49	50	51
E/E Eng.	247	63	63	64	64	64	64	61
Industrial Eng.	42	48	51	54	56	51	51	51
Materials Eng.	117	51	55	55	56	58	60	54
Mechanical Eng.	200	50	55	59	59	59	59	58
Nuclear Eng.	44	37	40	35	40	46	46	35
All Other Eng.	178	49	48	50	51	51	50	49
Total, All S&E	3,653	40%	42%	43%	43%	43%	44%	42%

employment. Adjustments were made to the tabulations received from the Social Security Administration, principally to account for the fact that some jobs are still not covered under social security and for the fact that only 90 percent of the population of interest had valid social security numbers. Readers with an interest in these adjustments should refer to the technical appendix for details.

The bottom line of Table 1-1 gives the best single measure of a foreign national stay rate. It indicates that 42 percent of the foreign nationals who received science and engineering degrees in 1984 were working in the United States in 1992.

Table 1-1 indicates that the stay rate for 1984 physical sciences graduates was 45 percent in 1992 and that it was within the range of 45 percent to 48 percent throughout the period from 1986 to 1992. The stay rate for engineering graduates was also stable over this period and was estimated at 55 percent in 1992. The stay rates for the social sciences and life sciences were substantially lower. Only in the life sciences did the stay rate increase substantially from the level observed in 1986. As these estimates are group statistics, they are by nature "net" estimates. That is, though the stay rate for engineering was 55 percent in both 1988 and in 1992, there was undoubtedly some movement abroad by members of this 1984 class who were working here in 1988. However, as the group statistics did not decline, such departures that occurred were offset by other members entering the U.S. workforce after residing abroad in 1988.

Table 1-1 also reports stay rates for subfields within engineering. When the number of graduates on a given line is less than about 250, we did not use a sample but requested the Social Security Administration to compute statistics for the whole group. Thus, the estimates provided for each subfield of engineering (and for all engineering) is free of sampling error, and we need not mistrust the estimates because of the small size of some of the groups. Since about 93 percent of all the engineering Ph.D.s in the Doctorate Records File had valid social security numbers, we have a high level of confidence in the resulting estimates. The fact that the highest stay rate was recorded for aeronautical engineers and the lowest for agricultural and biological engineers may reflect differences in the job opportunities available for engineering graduates in these fields in their home countries versus the United States. We can only speculate on the reasons for these differences, but the margin of error is small enough that we are certain that they are real differences.

Table 1-2 presents the 1992 estimates for Table 1-1 in another format. In Table 1-2, we describe the 1992 estimate from Table 1-1 as a mid-case estimate. We also present low and high estimates of the same stay rates. Because we had to make some assumptions regarding the number of persons employed in jobs not covered by social security and about the stay rate of persons who

Table 1-2. Temporary Residents Receiving S&E Ph.D.s in 1984 Who Were Working in the United States in 1992

	Percent Working in United States		
Degree Field	Low	Mid-Case	High
Physical Sciences	43%	45%	47%
Life Sciences	29	33	39
Social Sciences	21	26	30
Engineering	52	55	57
Aeronautical Eng.	60	64	67
Ag. & Bio. Eng.	21	23	24
Chemical Eng.	62	65	67
Civil Eng.	48	51	53
E/E Eng.	58	61	64
Industrial Eng.	48	51	54
Materials Eng.	51	54	56
Mechanical Eng.	56	58	60
Nuclear Eng.	33	35	37
All Other Eng.	46	49	51
Total, All S&E	39%	42%	45%

did not have social security numbers, the mid-case (though it is our best estimate), has a margin of error even where there was no sampling. Our low estimates for 1992 assume that all of the doctorate recipients missing a valid social security number left the United States after graduation and never came back. Our high estimates assume that persons missing social security numbers stayed to work in the United States at the same rate as persons with social security numbers. The mid-case estimate assumed that these persons stayed at half the rate of others with social security numbers in the same discipline.

We can be quite confident that the true stay rate for engineering is within the range shown in Table 1-2. In this discipline, there was no sampling done. In the others, however, we estimated the stay rate for a sample of 200 persons. The estimate of a stay rate of 45 percent for physical sciences could also have been affected by sampling error. We estimate the standard error of this estimated proportion to be 3 percent. Thus, at a 95 percent confidence level, we can estimate the (mid-case) stay rate for physical sciences to be in the range of 39 to 51 percent. We have not computed standard errors for all the estimates presented in Table 1-1, but note that physical

sciences had the lowest sampling rate (200 of 987) and thus would be expected to be most affected by sampling error. In most of the data presented below, sampling error is not an issue because the size of the groups was smaller than 200 and the behavior of the entire group was measured.

Table 1-3 shows stay rates defined in a manner similar to those in Table 1-1, but this time for persons completing a degree in 1987 or 1988. For these more recent graduates, the overall science and engineering 1992 stay rate is little changed, at 41 percent, however some disciplines have changed. Compared with the class of 1984, the main differences are that the engineering stay rate is lower and the social sciences stay rate is higher. However, the similarities between the class of 1984 and these later classes are greater than the differences. In each year, engineering and physical sciences show above-average stay rates. Also, the average for all science and engineering degree recipients is within the range of 41 to 44 percent if we wait 3 or more years past graduation to measure the stay rate. We observe slightly lower stay rates immediately after graduation, but we are not sure whether this is real or an artifact of the data. During the first 2 to 3 years after graduation, a high proportion of new Ph.D.s are working on postdoctoral appointments, and some of the persons receiving postdoctoral stipends appear not to pay social security taxes on the stipend. Thus, estimates of stay rates in years near the year of graduation may include some undercounting that we have not attempted to correct.

Table 1-3. Temporary Residents Receiving S&E Ph.D.s in 1987 or 1988 Who Were Working in the United States, by Degree Field, 1989-1992

Degree Field	Temp. Res. Ph.D.s	Percent Working in United States			
		1989	1990	1991	1992
Physical Sciences	2,838	36%	45%	45%	46%
Life Sciences	1,676	22	27	30	32
Social Sciences	1,574	28	30	30	30
Engineering	3,241	44	47	48	48
Total, All S&E	9,329	36%	40%	41%	41%

Table 1-4 provides detailed stay rates by broad field of degree and country of citizenship at the time of graduation, for 1987 and 1988 graduates who were temporary residents at the time of graduation. The countries shown provide about as much detail as was possible to obtain, given that we could not examine very small groups without the risk of violating confidentiality requirements of the Social Security Administration.

Table 1-4 indicates considerable variation in stay rates by country of citizenship. The country differences that appear for engineering in the first section of the table tend to hold up for the other disciplines as well. In engineering, graduates from three countries had stay rates well above average: India (77 percent), Iran (72 percent), and the Peoples Republic of China (66 percent). Countries with stay rates well below average include Japan (12 percent), Brazil (15 percent), and Korea (20 percent). All 6 of these countries exhibited a similar pattern in the other disciplines — insofar as the data was not suppressed for reasons of confidentiality.

When producing estimates for the life sciences and social sciences, some countries had to be combined for reasons of confidentiality. For example, we could not report on Brazil, Mexico, or Egypt separately in the life sciences because it would have violated the Social Security Administration's confidentiality rules. However, the stay rate of the Brazil/Mexico combined in the life sciences in only 13 percent and this is entirely consistent with a pattern of Brazil having stay rates substantially below average in all fields. The only real exception to the pattern noted is that one of the three countries with stay rates well above average, the Peoples Republic of China (PRC), had a stay rate not far above average in 1992 in the social sciences. Since fewer than 5 percent of the Ph.D. degrees earned by students from the PRC were in the social sciences, this does not change the general pattern that the PRC students stayed to work in the United States at a rate that is well above average.

The case of the PRC students is interesting in another respect. The 1987 and 1988 graduates from this country had a much higher stay rate recorded in 1992 than they did in 1989. Such a sharp increase in the stay rate from 1989 to 1992 was not observed for other countries. We may infer that a significant number of PRC citizens returned to the United States as a result of the political unrest and government repression that occurred in 1990.

Table 1-5 shows stay rates for 1990 temporary residents, but only for the broad degree fields of engineering and physical sciences. Table 1-5 shows separate stay rates for 4 groups identified by their response to questions on the Survey of Earned Doctorates. All 1990 Ph.D. recipients fall into one of these 4 categories. These results are of particular interest because these questions are asked of new Ph.D.s annually. If we could infer actual stay rates from graduating students' statements about their postgraduation plans, we could use the Survey of Earned Doctorates to

Table 1-4. Temporary Residents Receiving S&E Ph.D.s in 1987 or 1988 Who Were Working in the United States, by Degree Field and Country of Citizenship, 1989-1992

Degree Field and Country of Citizenship	Temp. Resident Ph.D.s	Percent Working in United States			
		1989	1990	1991	1992
Degree Field = Engineering					
Taiwan	648	51%	52%	54%	53%
Korea	457	25	21	22	20
Japan	35	9	9	9	12
Peoples Republic of China	200	35	55	60	66
India	393	72	79	77	77
Iran	89	65	68	72	72
Other Asia/Pacific	476	47	49	49	45
Egypt	81	17	20	22	20
Other Africa	135	32	37	41	45
Greece	98	52	50	48	47
Other Europe	165	36	42	42	38
Brazil	58	10	15	19	15
Mexico	40	39	51	53	51
Other Central/South America	81	30	35	39	41
Canada	46	47	47	47	47
Total, Engineering	3,002	44%	47%	48%	48%
Degree Field = Physical Sciences					
Taiwan	338	45%	50%	46%	48%
Korea	286	19	19	17	15
Japan	43	21	30	30	8
Peoples Republic of China	386	31	55	60	67
India	273	59	70	71	71
Iran	56	60	64	62	64
Other Asia/Pacific	396	44	47	49	49
Egypt	92	37	44	44	44
Other Africa	83	34	40	40	43
Greece	52	33	41	41	48
Other Europe	358	32	35	38	37
Brazil	46	18	21	12	12
Mexico	44	34	40	46	46
Other Central/South America	115	39	44	44	44
Canada	77	32	34	34	32
Total, Physical Sciences	2,585	38%	45%	45%	46%

Table 1-4. Temporary Residents Receiving S&E Ph.D.s in 1987 or 1988 Who Were Working in the United States, by Degree Field and Country of Citizenship, 1989-1992 (continued)

Degree Field and Country of Citizenship	Temp. Resident Ph.D.s	Percent Working in United States			
		1989	1990	1991	1992
Degree Field = Life Sciences					
Taiwan	190	35%	43%	37%	42%
Korea	138	23	19	18	20
Japan	19	21	30	30	8
Peoples Republic of China	106	23	42	52	65
India	103	46	56	64	66
Iran	18	35	35	53	47
Other Asia/Pacific	333	19	22	25	25
Africa	193	19	21	24	28
Greece	19	19	32	32	32
Other Europe	108	22	29	31	34
Brazil and Mexico	111	5	7	10	13
Other Central/South America	128	16	20	24	26
Canada	77	14	22	21	22
Total, Life Sciences	1,543	22%	27%	30%	32%
Degree Field = Social Sciences					
Taiwan	102	31%	32%	31%	27%
Korea	197	11	9	10	9
Japan	69	28	31	30	28
Peoples Republic of China	25	21	26	34	38
India	79	57	57	56	56
Iran	34	60	64	70	67
Other Asia/Pacific	353	24	27	28	28
Africa	179	30	31	31	31
Greece	21	30	35	35	35
Other Europe	171	38	42	41	41
Central/South America	121	17	19	19	19
Canada	89	26	30	32	32
Total, Social Sciences	1,440	28%	30%	30%	30%

Note: Totals for each degree field are smaller in this table than in Table 1-3 because this table excludes temporary residents with country of citizenship unspecified.

Table 1-5. Temporary Residents Receiving S&E Ph.D.s in 1990 Who Were Working in the United States, by Degree Field and Postgraduation Plan, 1989-1992

Degree Field and Postgraduation Plan	Temp. Res. Ph.D.s	Percent Working in United States			
		1989	1990	1991	1992
Engineering					
Postgraduation Plan:					
Definitely going abroad	446	3%	4%	8%	7%
Definite work in United States	456	44	86	94	94
Definite study in United States	193	11	41	61	65
No firm plans/no response	1,173	12	42	56	53
Total, Engineering	2,268	16%	44%	55%	54%
Physical Sciences					
Postgraduation Plan:					
Definitely going abroad	353	0%	8%	14%	17%
Definite work in United States	253	35	72	82	84
Definite study in United States	453	11	40	63	70
No firm plans/no response	877	7	24	40	44
Total, Physical Sciences	1,936	10%	31%	46%	51%

track changes in stay rate behavior. However, there are several pitfalls in any such effort. First, nearly half of the new Ph.D.s in these disciplines did not respond to that question or responded that they had no firm plans at the time of the survey. Second, it appears that only 65 to 70 percent of those with definite plans for postdoctoral appointments in the United States were working in the United States 2 years after graduation.[2] Third, behavior does not always conform with plans. Note that in each of these disciplines the group that reported definite plans to go abroad had a non-zero stay rate. In physical sciences the rate increased to 17 percent by 1992. A high proportion of physical sciences degrees were awarded to students from the PRC in 1990 (and a substantial share of engineering students as well). Thus, it is likely that the unrest in China in 1990 and the subsequent easing of immigration restrictions for PRC students are responsible for some of this.

One result from Table 1-5 that can be compared to results from earlier tables is that the overall stay rate for both engineering Ph.D.s and physical sciences Ph.D.s was more than 50 percent in 1992. Our analysis of the classes of 1987 and 1988 indicated a downturn in the engineering stay

rate compared with the class of 1984. However, when we compare the class of 1984 with the class of 1990, we observe little change in the stay rate for engineering (from 52 percent for 1984 graduates to 54 percent for 1990 graduates), while the physical sciences stay rates increases (from 46 percent for 1984 graduates to 51 percent for 1990 graduates) when these stay rates are computed 2 years after graduation. While there may be many factors at work, the high stay rate of PRC students and the increase in their numbers over this period are more than enough to explain these modest increases.

Table 1-6 shows stay rates for foreign nationals who were permanent residents at the time of graduation. Overall, the stay rates are much higher than stay rates for temporary residents. Our estimate for 1984 graduates in 1992 is 78 percent and 82 percent for 1988 graduates in 1992.

Permanent residents seem to be quite similar to U.S. citizens in that nearly all stay in the United States. However, there are some interesting similarities between the permanent residents

Table 1-6. Permanent Residents Receiving S&E Ph.D.s in 1984 or 1988 Who Were Working in the United States, 1989-1992

Degree Field	Perm. Res. Ph.D.s	Percent Working in United States			
		1989	1990	1991	1992
1984 Graduates					
Physical Sciences	197	90%	88%	87%	85%
Life Sciences	149	75	73	77	79
Social Sciences	214	67	66	63	63
Engineering, Total	272	91	89	85	84
Engineering, Taiwan	58	95	94	88	83
Engineering, Korea	20	51	46	41	41
Total, 1984 Graduates	832	82%	80%	78%	78%
1988 Graduates					
Physical Sciences	252	83%	83%	83%	83%
Life Sciences	262	66	70	74	80
Social Sciences	244	76	77	77	75
Engineering, Total	366	90	89	87	86
Engineering, Taiwan	71	96	92	92	87
Engineering, Korea	22	70	65	65	65
Total, 1988 Graduates	1,124	80%	81%	81%	82%

described in Table 1-6 and the temporary residents described in Tables 1-1 and 1-4. For each of these two categories of visa status at graduation, Ph.D. recipients in the physical sciences and engineering fields were the most likely to stay in the United States after graduation. Also, for each of the two visa categories, engineering students from Korea were more likely than all engineering students to leave the United States upon graduation. We did not produce statistics on permanent residents' stay rates beyond those in Table 1-6, partly because the number of permanent residents Ph.D. recipients is relatively small compared with the temporary resident category.

We can obtain an estimated stay rate for all Ph.D.s who were foreign citizens at the time they received their degrees combining the estimates for permanent residents and temporary residents. These are shown in Table 1-7. The overall stay rate of 49 percent for 1984 graduates in 1992 and 48 percent for more recent graduates in 1992 is probably the best rate to use when asking about the contribution of foreign-born scientists and engineers to the United States. However, when one is asking about the contribution of foreign students to the United States, it is usually more appropriate to use rates for persons who were temporary residents at the time of graduation, i.e. the rates shown in Tables 1-1 through 1-5. Most institutions define foreign students as students studying on temporary visas. Students on permanent resident visas have virtually all the rights of U.S. citizens in the employment arena. Perhaps the most important distinction, however, is that a substantial proportion of students on permanent visas at the time of graduation were on permanent visas at the time they first enrolled. That is, they immigrated first and then enrolled in school.

Table 1-7. Foreign Nationals Receiving S&E Ph.D.s in 1984, 1987, or 1988 Who Were Working in the United States in 1992

	Percent Working in United States	
Degree Field	1984 Graduates	1987 or 1988 Graduates
Physical Sciences	52%	52%
Life Sciences	42	44
Social Sciences	34	40
Engineering, Total	60%	55%
Total, All S&E	49%	48%

Chapter 2
STATISTICAL ANALYSIS OF FOREIGN NATIONAL RESPONDENTS TO THE SURVEY OF DOCTORATE RECIPIENTS

This chapter examines differences in characteristics and emigration decisions of 1981-1986 Ph.D. recipients known to have had temporary visas at the time of graduation. Data on certain characteristics and emigration for each individual in this group were observed from the year of graduation through 1991 for this analysis. This allowed a study of the behavior of those foreign students who worked in the United States from 5 to 10 years after graduation and then returned home, unlike many foreign students who returned home immediately after graduation. The purpose of this analysis was twofold. The first objective was to merely observe and then compare measurable characteristics of those among the group who reported in 1991 that they resided in the United States (referred to as stayers) with those who reported in 1991 that they resided outside the United States (referred to as leavers). The second objective was to develop a statistical model that could determine if the leavers were statistically different in terms of productivity in the labor force from the stayers.

DATA

The data set used for this analysis was the "1991 Survey of Doctorate Recipients Longitudinal File," which was generated from the longitudinal survey of science, engineering, and humanities doctorates funded by several federal agencies and conducted by the National Research Council. The survey began in 1973 and has been conducted every two years. The sample changes every two years as the oldest cohort is dropped and a cohort of new graduates is added. Until 1989, the sample included Ph.D.s who had graduated during the prior 42-year period. However, in 1991, the criterion for being dropped from the sample changed from Ph.D. year of more than 42 years past to age of 76 or older. The sampling frame for this survey is compiled from the Doctorate Records File, which is an ongoing census of all research doctorates earned in the United States.

The sample used in this chapter is a very small subset of the Ph.D. data file described above. Based on a valid response to the survey in 1987 and 1991, the sample was selected to include only those Ph.D. recipients from January 1, 1981 through June 30, 1986 whose degree field was either science or engineering (humanities were excluded), who were temporary residents at time of degree, and who were residing in the United States in 1987. This resulted in a sample of 264 persons of which 241 stayers were still residing in the United States in 1991, and 23 leavers were residing outside the United States.

COMPARISON OF MEASURABLE CHARACTERISTICS

The first objective of this analysis was to summarize and compare measurable characteristics for Ph.D. recipients from U.S. institutions who were on temporary visa status at the time of Ph.D. completion. Descriptive information was calculated for the group as a whole, and for the groups called leavers and stayers. Looking at these characteristics was the first step in establishing patterns among those types of foreign students who remained in the United States after degree completion compared with those who tended to work for a few years and then leave. The data for each of these groups are provided in Table 2-1 that follows.

Comparisons of annual salaries in 1987 and 1991 suggest that those who chose to leave the United States in 1991 were on average making lower salaries. Not only were average salaries in 1987 lower among those who did leave, salary growth for this group during the period was slower. At first, it appears that perhaps those foreign students who remained in the United States were more desirable to employers since they appear to have received higher salaries on average when compared to those who left after a few years of working. Caution should be taken with this salary comparison, however, because the 1991 average salary for leavers is based on 12 observations and converted from foreign currency to U.S. dollars.

Two key factors, choice of Ph.D. field and type of employment after graduation, could explain some of the observed average salary differences. Ph.D. field differed between the two groups. Leavers were more likely than stayers to have studied earth/environmental/marine sciences, and psychology/social sciences. They were less likely to have studied engineering. This distribution of degree fields could explain some of the difference in mean salary between the two groups. Regardless of productivity, market conditions cause salary differences across fields.

Employer types in 1987 were also different for stayers versus leavers. The largest difference in 1987 was in the percentage who were employed in tenured or tenure-track academic positions. Sixty-five percent of the leavers and 63 percent of the stayers were employed in academics in 1987, but the percentage without tenured or tenure-track positions was much higher for leavers versus stayers, 48 and 34 percent respectively. Over the four-year period from 1987 to 1991, stayers shifted out of nontenured into tenured positions as well as into business/industry. For the leavers, the percent employed in business/industry remained unchanged over the period, as did the percentage in tenure-track positions, but many of those in nontenured positions in 1987 shifted to government or nonprofit firms in 1991. Both the distribution across job types in 1987 and the change in this distribution in 1991 could explain differences in average salaries for leavers and stayers.

Table 2-1. Descriptive Information for Foreign National Respondents to the SDR

Characteristics	Combined Stayers and Leavers	Stayers	Leavers
N	264	241	23
Mean Salary in 1991	$55,300	$55,590	$42,800
Mean Salary in 1987	$36,600	$36,800	$35,100
Mean Age in 1991	38	38	37
Year of Ph.D. Degree:			
1981	12%	13%	0%
1982	13%	13%	13%
1983	15%	16%	9%
1984	16%	17%	9%
1985	25%	23%	43%
1986	19%	18%	26%
Year of Bachelor's Degree:			
Prior to 1974	25%	26%	10%
1974	9%	9%	10%
1975	12%	12%	19%
1976	12%	13%	9%
1977	12%	13%	0%
1978	11%	10%	14%
1979	9%	8%	19%
1980	9%	8%	19%
1981 or 1982	1%	1%	0%
Jones-Lindzey Reputational Rating of Ph.D. Department (mean)	3.30	3.27	3.73
Time to Ph.D. = Year of Ph.D. Minus Year of Bachelor's (mean years)	8.5	8.5	8.0
Females	27%	26%	30%
Married in 1991	76%	79%	43%
Married in 1987	63%	64%	57%
Have children in 1991	74%	77%	39%
Have children in 1987	42%	43%	30%
Ph.D. Fields:			
Math/Statistics/Computer Sciences	11%	11%	13%
Physics/Astronomy	8%	8%	4%
Chemistry	8%	8%	4%
Earth/Environmental/Marine Sciences	4%	3%	13%
Engineering	35%	36%	30%
Life Sciences	21%	22%	18%
Psychology/Social Sciences	13%	12%	18%
Employer Types in 1991:			
Business/Industry	31%	32%	22%
Academic—Tenured or Tenure-Track	37%	39%	17%
Academic—Nontenured	17%	16%	22%
Government or Nonprofit	15%	13%	39%
Employer Types in 1987:			
Business/Industry	25%	25%	22%
Academic—Tenured or Tenure-Track	28%	29%	17%
Academic—Nontenured	35%	34%	48%
Government or Nonprofit	12%	12%	13%
Participated in Postdoctoral Appointment in 1991	2%	2%	0%
Participated in Postdoctoral Appointment in 1987	18%	18%	22%
Participated in Postdoctoral Appointment in any year since graduation	24%	24%	22%
Country of Citizenship at Time of Degree:			
Europe	14%	13%	26%
Eastern Asia	35%	36%	23%
Western Asia	35%	35%	30%
Canada/Mexico	12%	12%	17%
Africa	4%	4%	4%

Salary differences might also be explained by the rate at which students in each group participated in postdoctoral appointments, which were more likely to have lower salaries. For both groups, approximately one in five students participated in a postdoctoral appointment in 1987.

While both choice of Ph.D. field and employment after graduation seemed to explain the average salary differences between these two groups, information on the quality rating of an individual's Ph.D.-granting department provided contrary information. Quality of Ph.D.-granting department was measured by a rating of the scholarly quality of department faculty calculated from the Jones-Lindzey Study in 1981 by the National Research Council. [See Jones, Lindzey and Coggeshall.] This study asked faculty members to rate departments on a scale from 0 (not sufficient for doctorate education) to 5 (distinguished) and a mean rating was then calculated for each department. By this measure, leavers attended departments which were, on average, of higher quality. The average Jones-Lindzey departmental rating for leavers was 3.73 versus 3.27 for stayers.

This quality measure was also calculated for all foreign students included in the "1991 Survey of Doctorate Recipients Longitudinal File" receiving degrees between 1981 and 1986 who were on temporary visas at the time of degree. This group was referred to as the total group of foreign students. This total group differs from our group of leavers and stayers in that a response to the survey in 1987 and 1991 was not required for this group. The mean Jones-Lindzey rating for this total group was 3.14. From this total group, a subset of people who were known to be in the United States in 1991 was generated. The mean Jones-Lindzey for this group of foreign students located in the United States in 1991 was 3.17. Since these two means, one for the group as a whole and one for the subset known to be in the United States in 1991, had almost the same average quality rating, it could be inferred that there was no superiority of the subset known to be in the United States when based on quality rankings of Ph.D. department. A mean Jones-Lindzey quality rating was also calculated for all U.S. citizens and permanent residents who received their Ph.D. between 1981 and 1986. This value was 3.19, again not significantly different from that for the total group of foreign students.

Comparison of demographic characteristics of the stayers and leavers provided information about the type of individual who was most likely to leave or stay after degree completion. The two groups were about the same age on average, but the leavers obtained more of their bachelor's and Ph.D. degrees in later years. While they might have obtained their degrees later in life, leavers completed them more rapidly. To measure the time required to obtain the Ph.D. degree, the difference between year of Ph.D. and year of bachelor's degree was calculated. For stayers, this was an average of 8.5 years versus an average of 8 years for leavers. Among the

leavers, the country of citizenship distribution was weighted more towards European or Canadian/Mexican descent as opposed to Asian descent. Only 53 percent of the leavers were Asian, which compares to 71 percent of the stayers. There was a slightly higher percentage of females among the group of leavers. Both in 1987 and 1991, leavers were less likely to be married or have children.

MODEL FOR MEASURING PRODUCTIVITY DIFFERENCES

Comparisons of these two subsets of foreign Ph.D. recipients provides insight into the value to the U.S. workforce of educating foreign science and engineering students. However, determining the value of education of foreign students necessitates a measure of productivity. If wage or salary information can be used as a measure of productivity in the workforce, then these data, as described above, can be used to analyze the value of educating foreign doctoral students. There are models in the literature that have been used to examine the value of educating foreign students. The brain drain theory asserts that the best students leave their country to get an education and never return. This can result from the fact that the country educating the student has a better idea of the student's true productivity than his home country; this is reflected in the wage offers. This model would imply that foreign students with above-average productivity stay in the United States. [See Kwok and Leland.] Another study by George Borjas [1989] found that earnings of immigrant scientists and engineers who eventually returned to their home country were about 11 percent lower than of those who stayed. However, this study examined persons who immigrated prior to 1970. These findings might not hold for more recent immigrants as substantial immigration law and regulation changes occurred during the late 1960s.

Table 2-1 indicates that average salaries of stayers were higher than those of leavers. If salaries can be used to measure productivity, then it might appear that stayers were the more productive group. However, observed salary differences between these two groups could result from differences between the two groups in other salary-determining characteristics as also measured and discussed above. Several factors, such as Ph.D. field and employment type, indicated that stayers would be expected to earn more than leavers, even if they were no more productive. Quality measures of department ranking provided contradictory information; while leavers earned less on average, they tended to come from higher-ranked departments. Thus, while the descriptive statistics derived from this data set provide interesting comparisons, they also provide contradictory evidence. From these measures alone, one cannot conclude that stayers were more productive simply because they had higher salaries on average. It is necessary to control for measurable differences between these two groups and take into account the relationship among these important explanatory variables.

In order to test the hypothesis that leavers were more productive than stayers as measured by salary differences, a simultaneous equations model was estimated to take into account the simultaneous determination of salary offers and the decision to emigrate, controlling for all measurable characteristics available in the data set. The model estimated included three equations, which allowed estimation of the effect of past productivity, salary in 1987, on the decision to leave the United States by 1991, as well as estimation of other factors affecting emigration. The final equation estimated current productivity, salary in 1991, controlling for self-selection into this sample, in order to measure the important influences of productivity among those remaining in the United States.

The primary hypothesis concerning productivity of leavers versus stayers can be tested with the first two equations in this model. The third equation relating to 1991 productivity for stayers was estimated to increase the understanding of determinants of productivity of foreign Ph.D.s who appear to have long-term plans to remain in the United States.

The model estimated is as follows:

Equation 1. ln(Salary in 1987) = f(gender, age, Ph.D. year, Jones-Lindzey Rating, Ph.D. field, employer type in 1987, postdoctoral participant in 1987, receive government support, country of citizenship)

This equation is estimated for leavers and stayers combined.

Equation 2. Probability of leaving in 1991 = f(salary in 1987, gender, age, marital status, children present, father's education, employer type, postdoctoral participant in any year, receive government support, country of citizenship)

The dependent variable in this equation is coded as a 0 or 1 depending on whether an individual stayed in or left the United States by 1991.

Equation 3. ln(Salary in 1991) = f(sex, age, Ph.D. year, Jones-Lindzey Rating, Ph.D. field, employer type in 1991, postdoctoral participant in 1991, receive government support, country of citizenship)

This equation is estimated for stayers, after controlling for self-selection into this 1991 salary sample.

In the above model, the decision of foreign persons educated in the United States to leave the country after several years of employment depended on the expected gain from such a location change. This expected gain depended on a comparison of opportunities in the United States to opportunities in the home country. Under the assumption that the wage measures the productivity of labor, opportunities in the United States can be reflected by salary offers in the recent past. Thus, a significant factor in predicting the probability that an individual left the United States should be salary in 1987. This salary offer was not exogenous to the decision to leave. Many of the same factors, both measurable and unmeasurable, such as motivation and culture, affected jointly the decision to leave and the 1987 salary offer. This implies that, in estimating Equation 2, it is not sufficient to use the 1987 salary variable as an explanatory variable since it is likely to be correlated with the error term.

A two-step procedure was used to estimate Equations 1 and 2. First, ordinary least squares estimates were used to generate predicted 1987 salary offers. In the second stage, the predicted 1987 salary was then used as an instrumental variable in the estimation of a probit model which estimated the probability that an individual leaves the country. The results of this two-step procedure are shown in Tables 2-2 and 2-3.

As Table 2-2 shows, the primary factors that affected 1987 salary offers were Ph.D. year, degree field, and type of employer. These variables had the expected signs. For the degree field variable, the omitted category was engineering; thus, relative to engineers, persons in every other field classification, with the exception of persons in math/statistics/computer sciences, earned lower salaries. For the employer type variable, the omitted category was government or nonprofit firm. Relative to this omitted group, persons employed in business/industry earned more and persons employed in nontenured academic positions earned less. As expected, those graduates who were employed in a postdoctoral position in 1987 earned substantially lower salaries. Factors which appeared to have no explanatory power included gender, age, the Jones-Lindzey department reputational rating, and country of citizenship. Predicted 1987 salary was generated from this equation and used as an instrumental variable in Equation 2, which estimated the probability of leaving the country.

Equation 2 was estimated using a probit model. The dependent variable in this model was equal to one if an individual left the country and equal to zero otherwise. This model assumed that this probability of leaving was determined by the factors listed above in Equation 2. These results are presented in Table 2-3. The most important result here is the insignificance of the predicted 1987 salary variable. One of the two primary objectives of this chapter was to test whether leavers were statistically different in terms of productivity. Assuming that productivity is

reflected in salary offers, finding that the predicted value of the 1987 salary offer was significantly different from zero in terms of explaining the probability of leaving would satisfy this objective. The t-ratio on this variable indicated that we could not reject the hypothesis that the effect of this variable was zero.

Table 2-2. OLS Regression of 1987 Salary Equation

Dependent Variable: ln (1987 Salary)
Dependent Variable Mean: 10.45

Explanatory Variable	Coefficient Estimate	t-ratio*
Constant	12.5600	14.89*
Gender (Female = 1)	-0.0068	-0.21
Age	-0.0032	-0.82
Jones-Lindzey Reputational Rating (Dummy = 1 if average rating for department is $\geq$ 3.5)	-0.0045	-0.14
Ph.D. Year	-0.0217	-2.29*
Ph.D. Field Dummy Variables:**		
Math/Statistics/Computer Sciences	-0.0635	-1.25
Physics/Astronomy	-0.1306	-2.21*
Chemistry	-0.2392	-3.97*
Earth/Environmental/Marine Sciences	-0.2061	-2.81*
Life Sciences	-0.1942	-4.60*
Psychology/Social Sciences	-0.1085	-2.19*
Employer Type Dummy Variables:**		
Business/Industry	0.1543	2.72*
Academic -- Tenured or Tenure-Track	0.0558	1.04
Academic -- Nontenured	-0.1958	-3.75*
Receives Government Support Dummy	-0.0033	-0.10
Postdoc Participation (in 1987) Dummy	-0.2963	-6.43*
Country of Citizenship Dummy Variables:**		
Europe	0.0076	0.17
Western Asia	-0.0158	-0.46
Canada/Mexico	0.0562	1.18
Africa	-0.0833	-1.15

Adjusted R^2 = 0.6380
F Value = 23.635
Standard Error of Estimate = 0.2079
N = 245

* Significant at α = .05 level
** For Ph.D. Field, engineering is the excluded category.
For Employer Type, government/nonprofit is the excluded category.
For Country of Citizenship, Eastern Asia is the excluded category.

Table 2-3. Probit Model Estimating the Probability of Leaving

Dependent Variable = 1 if leaver
= 0 if stayer
Dependent Variable Mean: 0.09

Explanatory Variable	Coefficient Estimate	t-ratio*	Marginal Effects (percentage points)
Constant	15.127	1.75	
1987 Predicted Salary	-1.171	-1.52	-1
Gender (Female = 1)	-0.342	-1.08	-3
Age	-0.068	-1.67	-0
Married Dummy	-0.709	-1.83	-6
Children Present Dummy	-0.175	-0.47	-1
Father's Education Dummy (College Graduate = 1)	0.458	1.54	4
Employer Type Dummy Variables:**			
Business/Industry	-1.315	-3.18*	-9
Academic – Tenured or Tenure-Track	-1.163	-3.00*	-9
Academic – Nontenured	-1.022	-2.31*	-6
Receives Government Support Dummy	-0.753	-2.25*	-7
Postdoc Participation (in any year) Dummy	-0.709	-1.46	-6
Country of Citizenship Dummy Variables:**			
Europe	0.114	0.28	1
Western Asia	-0.196	-0.58	-1
Canada/Mexico	0.152	0.34	1
Africa	-0.161	-0.21	-1

Log-Likelihood = -57.92176
N = 264

* Significant at $\alpha = .05$ level
For Employer Type, government/nonprofit is the excluded category.
For Country of Citizenship, Eastern Asia is the excluded category.

The only explanatory variables that were significant were the employment type variables and the government support variable. They all had the expected signs. A foreign national Ph.D. was less likely to leave the country if employed in any of the employer types relative to the omitted category, which was again government or nonprofit firms. He was also less likely to leave if he was receiving government support in his work.

Determining the influence of each explanatory variable on the probability of leaving using a probit model requires calculating the marginal effects. These values are shown in the last column of Table 2-3. Marginal effects measure the change in the probability of leaving resulting from a one unit change in an explanatory variable, holding all others constant; the marginal

effects are calculated at the mean values of the independent variables.[6] For the 1987 predicted salary variable, the unit of change was $4000 annually. All other variables are dummy variables and the change was from a value of one to a value of zero. Thus while the 1987 predicted salary variable is not statistically significant in this particular sample, a one percent decrease in the probability of leaving results when annual salary is increased by $4000. While this is a small change in probability, it should be compared with an overall probability of leaving of only 9 percent, as shown by the dependent variable mean. Those variables which were statistically significant did impact the probability of leaving even more. Being employed in business or industry, relative to government/nonprofit employment, reduces the probability of leaving by 9 percentage points.

Equations 1 and 2, which were used to test one of the primary objectives of this research, were estimated with a sample of foreign students responding to the survey in both 1987 and 1991. This resulted in a total of 264 students, 23 of whom left the country by 1991, or approximately 9 percent of the sample. This sample of definite responders in both years can be referred to as the pure sample. In order to determine whether the insignificance of past productivity (predicted 1987 salary) in the decision to leave the country merely resulted from the small number of leavers, the same model was tested with what can be called an extended sample. This extended sample utilized information from the 1989 survey for individuals who responded in 1987 but not in 1991. Individuals who either reported a foreign location in 1989 or had surveys mailed to them in 1989 at a current address that was foreign and deliverable were added to our group of leavers. This increased the number of leavers from 23 in the pure sample to 64 in the extended sample. Equations 1 and 2 were estimated again, and the predicted 1987 salary offer was again not significantly different from zero in predicting the probability of leaving the country.

Finally, 1991 productivity for stayers, Equation 3, was estimated to increase the understanding of determinants of productivity of foreign Ph.D.s who appeared to have long-term plans to remain in the United States. This equation was first estimated using a two-step estimation procedure which utilized predicted probabilities of leaving estimated from Equation 2 to correct for self-selection into the 1991 salary sample. These predicted probabilities were used to calculate the inverse Mill's ratio (λ) which was included as an explanatory variable in the 1991 salary equation. The estimated coefficient on λ indicated the direction and magnitude of the selectivity bias (see Heckman, 1979). Selectivity bias can result from the nonrandom selection of individuals into the stayer group. This procedure produced a coefficient on λ that was not significantly different from zero at an alpha level of .05 and, therefore, the estimates of Equation 3 shown in Table 2-4 resulted from an ordinary least squares regression.

Table 2-4. OLS Regression of 1991 Salary Equation

Dependent Variable: ln (1991 Salary)
Dependent Variable Mean: 10.88

Explanatory Variable	Coefficient Estimate	t-ratio*
Constant	14.3923	15.02*
Gender (Female = 1)	0.0019	0.04
Age	-0.0193	-4.23*
Jones-Lindzey Reputational Rating (Dummy = 1 if average rating for department is ≥ 3.5)	0.0781	1.99*
Ph.D. Year	-0.0332	-3.11*
Ph.D. Field Dummy Variables:**		
Math/Statistics/Computer Sciences	-0.0564	-0.95
Physics/Astronomy	-0.1533	-2.33*
Chemistry	-0.0748	-1.01
Earth/Environmental/Marine Sciences	-0.1075	-1.16
Life Sciences	-0.0663	-1.25
Psychology/Social Sciences	-0.0568	-0.94
Employer Type Dummy Variables:**		
Business/Industry	0.1742	2.60*
Academic – Tenured or Tenure-Track	-0.0112	-0.18
Academic – Nontenured	-0.3978	-5.36*
Receives Government Support Dummy	0.0874	2.23*
Postdoc Participation (in 1991) Dummy	-0.3037	-2.67*
Country of Citizenship Dummy Variables:**		
Europe	-0.0146	-0.28
Western Asia	0.0046	0.12
Canada/Mexico	0.1231	2.11*
Africa	0.0580	0.65

Adjusted R^2 = 0.4794
F Value = 11.761
Standard Error of Estimate = 0.2414
N = 223

* Significant at α = .05 level
** For Ph.D. Field, engineering is the excluded category.
For Employer Type, government/nonprofit is the excluded category.
For Country of Citizenship, Eastern Asia is the excluded category.

The significant factors in explaining the 1991 salaries of those in the stayer group included age, quality of Ph.D. department rating, Ph.D. year, Ph.D. field of physics/astronomy, employer type of business/industry or academic (nontenured), receiving government support, participation in a postdoctoral appointment in 1991, and a border country as country of citizenship. The negative sign on the age variable is uncommon as age is generally thought to capture work experience in many human capital models. In this sample of Ph.D.s, it is more likely that relevant work

experience was captured by the Ph.D. year variable instead, which has the expected sign. The simple correlation between age and Ph.D. year was not strong at -0.32. In this model, given the fact the Ph.D. year was restricted to a five-year period, age was possibly capturing the effect of obtaining a Ph.D. later in life. To the extent that these older Ph.D.s were less marketable or that they were older because they took longer to obtain their degree and were consequently less productive than those who finished quickly, a negative coefficient on the age variable would be expected.

CONCLUSION

This chapter has provided comparisons of characteristics and emigration decisions among a sample of foreign students earning U.S. doctoral degrees between 1981 and 1986. Descriptive information compared the group who was still residing in the United States in 1991, the stayers, to the group who was known to have emigrated by that year, the leavers. This descriptive information indicated that leavers appeared to have earned lower average salaries in 1987 and to have slower salary growth during a four-year period. Leavers were also less likely to have received a degree in engineering and more likely to have received a degree in the earth/ environmental/marine sciences or in psychology/social sciences. Leavers, on average, came from higher-ranked departments. A higher percentage of leavers were employed in academic positions that were nontenure-track in 1987 and in government or nonprofit organizations in 1991. Fewer leavers were married or had children, and this was true in 1987 as well as 1991.

This chapter has also examined the effect of past productivity on the decision to emigrate. A simultaneous system of equations was estimated to capture the effect of predicted 1987 salary offers on the decision to emigrate by 1991. This predicted salary variable was found to be insignificant in predicting the probability that an individual does emigrate. Thus, while a comparison of the average salaries for leavers and stayers may lead to the conclusion that leavers are the less productive group, the model estimated in this chapter refutes this finding. The model did however, confirm that foreign Ph.D. students employed in business/industry or academic jobs are less likely to emigrate compared with those employed by government or nonprofit firms.

TECHNICAL APPENDIX

This appendix provides detail about data and methods used to produce the results that are described in Chapter 1 of this report.

SOURCES OF DATA

The data for Chapter 1 consist almost exclusively of a set of more than 100 groups of Ph.D. recipients who received degrees from U.S. universities in 1984, 1987, 1988, and 1990. We first discussed this project carefully with staff of the National Research Council, the National Science Foundation, and the Social Security Administration to make sure that the methods chosen would comply with each organizations' policy regarding the confidentiality of data regarding individuals.

Our method was to start with responses to the Survey of Earned Doctorates for the years of interest. This survey is not a sample survey but rather a complete census of new doctorate recipients in the United States, administered at or near the time that they complete their doctorate. Among the questions asked of these persons are country of citizenship, degree field, and postgraduation plans. We used answers to these questions to define and identify groups for which we wanted to estimate a stay rate (e.g., temporary residents graduating in 1984 with a degree in engineering). The National Research Council then prepared a diskette containing the birth years and social security numbers of the persons in each of these groups. In most cases, we included all the persons with the traits used to define the group. However, a sample of 200 was used in cases where the total in the group was greater than about 250 persons. In total, we identified groups of foreign citizens containing a total of 11,219 persons. In addition, we specified 12 groups of U.S. citizens containing a total of 2,400 persons. These were used to help us make adjustments to the raw data received from the Social Security Administration.

The Social Security Administration first checked to identify persons for whom the social security numbers we provided were invalid. Also, they compared the year of birth we provided for each social security number with the year of birth in the social security files for the person with that number. They then excluded, from any tabulations, persons with invalid numbers and persons for whom the birth years differed by more than 1 year. The primary concern that led us to this birth year screen was the possibility that some social security numbers reported on the Survey of Earned Doctorates might be incorrect yet would be treated by the Social Security Administration as valid if it was identical to one of the millions of numbers in their system. By requiring the birth year to match or be off by no more than one year, we probably eliminated more than

95 percent of any such false matches. We found that 3.5 percent of U.S. citizens and 5.7 percent of foreign citizens had birth years that did not match within one year. This is not surprising since neither organization has 100 percent accuracy recording birth year. As far as the difference between the United States and foreign citizens in this regard, we postulate two distinct reasons. One is that foreign citizens sometimes write numbers differently or interpret questions differently. Another is that some foreign citizens do not have a social security number but may have reported a similar number issued by their university to students who don't have and don't want to get social security numbers. Insofar as the second reason holds, the difference between the U.S. rate of false matches (3.5 percent) and the foreign rate (5.7 percent) could be used as an indication of false matches that made it through the screen. That is, persons whose birth dates matched or were off by no more than one year were treated as having valid social security numbers. Since there are about 90 possible birth years (e.g., 1901 to 1991) that describe nearly all persons in the social security system in 1992, we conclude that a social-security-like number or a fake social security number would make it through our birth year screen with a chance of only about 3 out of 90, i.e., a probability of less than 3 percent. However, the chance of an invalid number making it through our screen is less than this. Many 9-digit numbers do not match because that number has not yet been issued to a person as a unique social security number. We did not carry this further after concluding that the possibility of false matches in our sample is quite small, surely less than 3 percent.

After screening out invalid social security numbers and numbers without a birth year that matched (or were off by no more than one year), the Social Security Administration made an initial set of computer tabulations by calculating for each group the proportion with earnings of $5,000 or more in each of the years from 1986 to 1992. This produced several groups where problems of confidentiality occurred. The practical application of the Social Security Administrations confidentiality rules meant that they would report no proportion if the group had a calculated proportion of 100 percent or 0 percent as this would permit the identification of individuals by persons who could match the social security numbers with names (e.g., the National Research Council staff who prepared the groups sent to the Social Security Administration). Further, to be safe, the Social Security Administration would not calculate a proportion if all but three persons in a group had earnings of $5,000 or more. Thus, for example, we produced separate estimates for 1987 and 1988 engineering and physical sciences graduates from Mexico and Brazil. However, we were forced to combine Mexico and Brazil with other Central and South American countries to produce stay rates in social sciences/psychology, and we were forced to combine Brazil and Mexico for life sciences stay rates. For Japan, we could not produce separate estimates for physical sciences and life sciences because of the confidentiality restriction. In this case, we chose to combine the two disciplines, resulting in identical stay rates for physical sciences and life sciences in Table 1-4. Combining separate

groups of disciplines and reporting out the combined rate for each discipline was not judged to be deceptive because the combined stay rate in 1992 was only 8 percent and represented only 5 persons who stayed out of a combined total of 62. The true rate for the physical sciences group would be under 12 percent even if all 5 persons were physical scientists.

ADJUSTMENTS TO DATA: MISSING AND INVALID SOCIAL SECURITY NUMBERS

One reason for missing or invalid social security numbers is data error. The respondent to the Survey of Earned Doctorates might fail to write down his or her number, or record his or her number incorrectly, or the coder at the National Research Council might make an error. If we were confident that other reasons were of no importance, we would not make any adjustments to account for missing social security numbers. However, we believe that sometimes social security numbers were missing because the foreign graduate did not have a social security number. The vast majority do have social security numbers, and this is not surprising as these are used by both banks and universities for identification numbers. However, it is possible for a student to go through graduate school without a social security number. Most universities will issue a similar 9-digit ID number to foreign students if they don't want to get a U.S. social security number. These usually start with the number 9, a number which the Social Security Administration never uses for the first digit of a true social security number. Many of our invalid social security numbers started with a 9, so it appears the student was confused and thought it was a social security number. But we also had a significant number of graduates for whom no social security number was recorded by the National Research Council. Among 1987 and 1988 graduates, the proportion with no social security number ranged from a low of 6.5 percent of engineering graduates to 16.4 percent of social sciences/psychology graduates. Further, there was variation by country: countries with the lowest proportion missing social security numbers tended to be countries with high stay rates, and countries with the highest proportion missing tended to be countries with low stay rates. This suggests that a substantial number of persons did not get social security numbers because they did not intend to work in the United States after graduation.

We made a low case assumption that all of the persons with missing or invalid social security numbers left the United States after graduation and did not return to the United States in subsequent years. However, this is obviously extreme. At the other extreme, our high case assumption was that the persons with missing or invalid social security numbers stayed to work in the United States at the same rate as others with the same characteristics (year of graduation, degree field, country of citizenship). Our mid-case estimate is always the average of the high and low cases. Thus, in the mid-case estimates, we are assuming that the stay rate for those with missing numbers is half the stay rate for those with valid social security numbers in the same

group. It turns out that the mid-case is not very different from the more extreme low case as the groups with a relatively high proportion of missing social security numbers (e.g., social sciences/ psychology graduates from Canada and from Central/South America) are almost always groups for which the holders of social security numbers recorded a stay rate of less than 35 percent.

OTHER ADJUSTMENTS

Other adjustments were needed because our control groups of U.S. citizens were recorded as earning $5,000 or more only about 90 percent of the time. To the extent this is so because they are working in jobs not covered by social security or because they were out of work for most (but not all) of the year, then the social security counts underestimate the number who are in the U.S. workforce. A few are missing because the person has died, is institutionalized, or is out of the labor force altogether. Of course, some U.S. citizens do leave the United States for foreign appointments, but our government collects no statistics on such movements. We developed the best data we could on each of these possible reasons for U.S. citizens not reporting social security covered wages.

We estimated death rates of U.S. citizens by using the age-specific death rates recorded by the TIAA insurance company.[3] We found out that state government employment in three states (Ohio, Massachusetts, and Alaska) is still completely exempt from social security taxes, including the persons employed by state universities in those states.[4] We used the Survey of Doctorate Recipients (SDR) to estimate the proportion of recent Ph.D.s in the sciences and engineering who were employed at universities in those states. We used Department of Education data to estimate the proportion among those employed in education in these states who were employed in public universities as opposed to private.[5] We also used 1991 SDR data to estimate unemployment rates, the percent who were out of the labor force, and the percent who were employed part-time. We assumed that at least one-half of one percent of U.S. citizens are working outside the United States (after examining the proportion who reported to the SDR that they had worked abroad for at least 3 months, and the proportion who reported that they were working abroad at the time of the 1991 SDR).

We made assumptions based on the best data available for all these factors and ended up with a set of adjustment factors that we judged would account for the U.S. citizens who were working in the United States or looking for work in the United States. These adjustment factors were then applied to the estimates of stay rates for foreign graduates as they too take employment in state government jobs that are not covered by social security, and they experience similar rates for death and unemployment or part-time employment. Doing this caused us to increase the rates for foreign national recipients of engineering degrees by a factor of 1.047 in the mid-case. The

adjustment factor for other disciplines in the mid-case were: 1.039 for physical sciences, 1.069 for life sciences, and 1.083 for social sciences. The higher adjustment factors for life sciences and social sciences/psychology were due to the fact that, compared with engineering and physical sciences, these fields had higher unemployment and part-time employment, and also a higher proportion of the recent graduates were employed in universities located in Ohio, Massachusetts, or Alaska.

The adjustments for missing social security numbers lowered the raw rates we obtained from the Social Security Administration. The other adjustments (e.g., for persons employed in non-covered employment) raised the rates. The net effect was that most stay rates reported in the main body of this report are very close to the unadjusted rates we obtained from the Social Security Administration when we asked for the proportion earning $5,000 or more. For example, for the 1984 graduates, we reported the following 1992 stay rates in Table 1-1: 55 percent in engineering (compared with an unadjusted rate of 55 percent); 45 percent in physical sciences (compared with 46 percent unadjusted); 33 percent in life sciences (compared with 35 percent unadjusted); and 26 percent in social sciences (compared with 28 percent unadjusted).

ENDNOTES

1. For all tables in Chapter 1, physical sciences includes mathematics, computer sciences, and physical sciences; social sciences includes psychology and social sciences.

2. Anecdotal reports and the data shown here for 1991 and 1992 indicate that some postdocs do not pay social security taxes. We made an adjustment for employment in jobs that aren't covered by social security (e.g., state universities in Ohio) for all of our estimates and are confident that this is appropriate for the totals in each discipline. However, this general adjustment may not be enough for those who did postdocs since foreign nationals take postdocs more frequently than U.S. citizens in the same discipline.

3. These were published in National Research Council, Biomedical and Behavioral Research Scientists: Their Training and Supply, Volume I, Washington, DC: National Academy Press, 1989, p. 114.

4. Telephone conversation with Mr. Ken Sanders of the Social Security Administration, July 1994.

5. National Center for Education Statistics, State Higher Education Profiles, Washington, DC, 1993 (NCES 93-169).

6. The marginal effect of explanatory value x_i on y, the probability of leaving, is $dy/dx_i = \beta_i f(x'\beta)$ where β_i is the coefficient on x_i reported in Table 2-3, $f(x'\beta)$ is the standard normal density, $x'\beta = \Sigma_{i=1}^{k} \bar{x}_i \beta_i$, and $\bar{x}_i$ is the mean value of each of the k explanatory values.

REFERENCES

Borjas, George J., "Immigrant and Emigrant Earnings: A Longitudinal Study," Economic Inquiry, January 1989, vol. XXVII, pp. 21-37.

Heckman, James J., "Sample Selection Bias as a Specification Error," Econometrica, January 1979, Vol. 47, No. 1, pp. 153-161.

Kwok, Viem and Hayne Leland, "An Economic Model of the Brain Drain," American Economic Review, March 1982, vol. 72, no. 1, pp. 91-100.

Jones, Lyle V., Gardner Lindzey and Porter E. Coggeshall, (eds.), An Assessment of Research-Doctorate Programs in the United States: Engineering, 1982, Washington: National Academy Press.