

This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Science and Engineering Careers in the United States: An Analysis of Markets and Employment

Volume Author/Editor: Richard B. Freeman and Daniel L. Goroff, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-26189-1

Volume URL: <http://www.nber.org/books/free09-1>

Conference Date: October 19-20, 2005

Publication Date: June 2009

Chapter Title: Supporting "The Best and Brightest" in Science and Engineering: NSF Graduate Research Fellowships

Chapter Author: Richard B. Freeman, Tanwin Chang, Hanley Chiang

Chapter URL: <http://www.nber.org/chapters/c11617>

Chapter pages in book: (p. 19 - 57)

---

# Supporting “The Best and Brightest” in Science and Engineering NSF Graduate Research Fellowships

Richard B. Freeman, Tanwin Chang, and  
Hanley Chiang

---

Stipends—payments to students as part of their studies—are a major form of income for graduate students in science and engineering (S&E) and a potential determinant of decisions to undertake graduate study.<sup>1</sup> Broadly defined to include fellowships, research assistantships, teaching assistantships, and postdoc awards, stipends account for upwards of one-fourth of the lifetime incomes of S&E Ph.D.s, depending on field, time to degree, and prevalence and length of postdocs.<sup>2</sup> In addition, fellowships are potentially the most attractive stipend because they signal to students that the granting institution views them as top prospects for success in graduate study and is willing to support their studies rather than requiring them to work

Richard B. Freeman holds the Herbert Ascherman Chair in Economics at Harvard University, and is the director of the labor studies program at the National Bureau of Economic Research. Tanwin Chang is an optical scientist at Adaptive Optics Associates. Hanley Chiang is a researcher at Mathematica Policy Research, Inc.

Supported by a grant from the NSF in collaboration with the Council of Graduate Schools (CGS). While serving as an AAAS/NSF Science and Technology Fellow, Dan Stanzione was instrumental in facilitating this study. Jason Abaluck of Harvard assisted in preparation of the data. We thank Myles Boylan of the NSF for a thoughtful review of a short version of this work published in the AEA proceedings.

1. In the 1990s, about two-thirds of S&E grad students received stipend support of some form and thus did not have to rely on loans, personal finances, or money from family and friends as their primary means of support (National Science Board 2004).

2. Fellowships pay much less than faculty jobs, but the fact that they are earned early in the career greatly augments their value relative to earnings from full-time employment. Using a five percent discount rate, a \$30,000 fellowship received at the first year of graduate study is comparable to \$49,000 received ten years later, \$79,500 received twenty years later, and \$129,300 received thirty years later. The discounting is such that scientists who did ten years of study and postdoctoral work and then worked for thirty years would earn 29 percent of their lifetime income during their fellowship years if their annual salary was twice that of their fellowship and would earn 21 percent of their lifetime income during their fellowship years if their annual salary was three times that of their fellowship.

to cover graduate school expenses. In 2001 13 percent of full-time graduate students reported that a fellowship or traineeship was their primary source of graduate support (National Science Foundation 2001a). Slightly more than one-third of those, or approximately 4 percent of full-time graduate students, obtained awards from federal government sources (National Science Board 2004).

Since its establishment in 1952, the National Science Foundation's (NSF) Graduate Research Fellowship (GRF) has been the United State's premier award for science and engineering graduate students. The NSF's stated purpose for the Graduate Research Fellowship is to "ensure the vitality of the scientific and technological workforce in the United States and to reinforce its diversity." Because NSF fellowships are limited to citizens or permanent resident aliens, they are a potentially important policy tool for inducing citizens and residents, including those from underrepresented groups, to study S&E and thus to raise their representation in the S&E workforce.

What determines the likelihood that an applicant wins a GRF award? Has the composition of applicants and awardees changed over time? What is the supply response of students to increases in the number or the amount of fellowships? How might alternative fellowship policies affect the numbers and measured attributes of awardees? Can we generalize from the response of students to GRFs to the possible impact of stipends on the supply of students to S&E more broadly?

This chapter uses data from the NSF's 1952 to 1993 Cumulative Index (CI) file, its ensuing updates, and other sources to examine these questions. The Cumulative Index plus updated data provides information on the over 200,000 individuals who applied for NSF graduate fellowships from 1952 through 2004, including the student's graduate record exam scores (GREs), grade point average (GPA), ratings of reference letters, (and from 1976 on), the ratings of applicants by review panels. We supplement these data with information on the earnings of college graduates, unemployment rates, and the sizes of graduating bachelor's cohorts. To determine the effects of changes in the number of GRFs and the amount of awards on the number of applicants and quality of awardees, we use time series fluctuations in the awards as the exogenous variation necessary to identify student supply responses.<sup>3</sup>

Section 1.1 describes the GRF program and how it has changed over

3. The 2002 report on the GRF by WestEd looks at what happens to fellows rather than what attracts students to the fellowships. It shows that in the time period leading up to 1993, around 70 percent of fellows received their baccalaureate from RUI institutions (as defined by the Carnegie Classification), and more than 90 percent went on to perform graduate work at RUI institutions. Some programs receive a disproportionately high number of fellows: 34 percent of Life Sciences Fellows enrolled in five institutions. The report shows that eleven years after beginning their Ph.D. program NSF Fellows had completion rates of around 70 percent, which exceeds the overall completion rate for entering students; that fellows pursued

time. It finds that the ratio of GRFs to S&E bachelor’s degrees fell by nearly 75 percent from the 1950s to the 1990s/2000s; that the distribution of awards shifted from physics and chemistry to biological sciences, social sciences, and engineering, and that women and minorities have increased their share of awards.

Section 1.2 uses a linear probability model to estimate the characteristics that help an applicant win a GRF. It finds that winning depends greatly on GREs, GPAs, and reference letters operating through panel ratings, and that, consistent with NSF’s diversity goal, women and minorities have higher chances of winning an award than white men with similar attributes.

Section 1.3 shows that an increase in stipend value increases the number of applicants, including those with very high skills, and thus raises the average skill of awardees. By contrast, increases in the number of awards are associated with modest declines in the average quality of awardees.

Section 1.4 explores new ways to determine the number and value of GRF awards, such as increasing the number and/or value of awards; indexing the number and value to measures of the supply of potential applicants and to the monetary attraction of alternative careers; and giving awards by the measured academic and personal characteristics of applicants, irrespective of field.

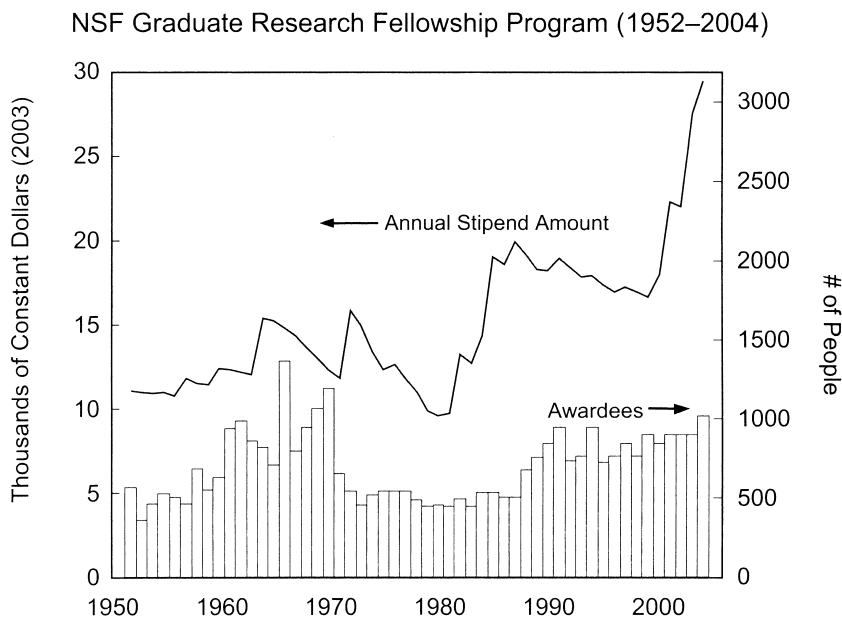
## 1.1 The GRF Program Over Time

Congress established the National Science Foundation’s Graduate Research Fellowship program in 1952 to support the vitality of the scientific and technological workforce in the United States. The program was open to U.S. citizens only. It gave a \$1,600 stipend and covered normal tuition and fees at the institution of choice. The NSF typically supported students for three years but this was not automatic; students had to reapply each year. The GRF has changed since its inception.<sup>4</sup> In 1963 it capped the cost-of-education, and in 1972 students no longer had to reapply (Goldsmith, Presley, and Cooley 2002). Today awardees may choose three years of support over a five-year period. They must be pursuing a research-based graduate degree in an NSF-supported field. The fellowship may be used outside the United States and is portable across institutions. In 1997 the NSF introduced new selection criteria that put greater weight than in the past in its awarding fellowships to “broader impacts” such as integration of research and education, diversity, benefit to society, and enhancement of scientific and technical understanding.

---

academic careers more than their peers, and that a slightly larger percentage reported having achieved traditional measures of accomplishment during graduate school (such as presentations, refereed articles, edited book parts, etc.).

4. National Science Foundation, Graduate Research Fellowship Program (2007) gives the details of the program as of 2007. The details change yearly with NSF policies.

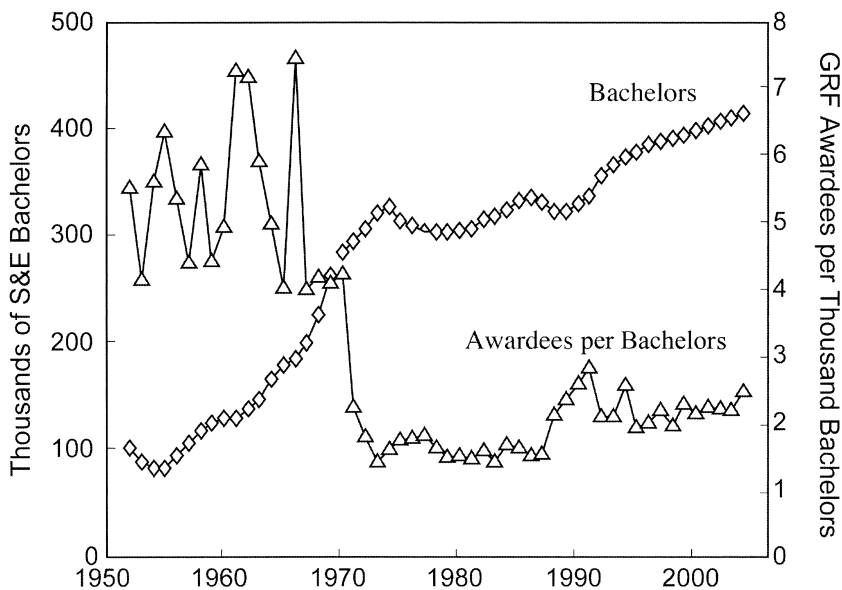


**Fig. 1.1** The number and value of GRF awards vary over time.

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related datasets.

The number of awards and the dollar value of stipends have changed over time with NSF budgets and decisions. Figure 1.1 shows the absolute number of awards and constant dollar value of the stipends from 1952 to 2004. In the 1950s the NSF gave around 500 to 600 awards. In the 1960s it gave around 1,000 awards, with a peak of 1,373 in 1966. In the 1970s and 1980s the number of awards fell to around 500, then rose in the late 1980s/early 1990s toward 900 to 1,000. The big jumps in the value of awards reflect NSF decisions to raise nominal stipend amounts. The NSF raised the value of fellowships in the early 1980s to the late 1980s and again in 1999 to 2004.<sup>5</sup> In 1999, the Committee of Visitor's (COV) report evaluating the program noted that "the GRF awards are no longer as attractive as they once were" (Committee of Visitors 1999). At \$15,000 per year for three years the awards were significantly lower than comparable fellowships. The COV speculated that this could (at least partially) explain the steady decline in applicants in the years leading up to 1999 and recommended that the stipend value be raised to \$18,000. In the subsequent four years, the value of the stipend rose 83 percent to \$27,500, which prompted the next COV (convened in 2003) to worry that the stipend was now too large.

5. The gradual declines in the stipend values in years when the NSF has not raised them reflect the impact of inflation. The NSF has never reduced the nominal value of awards.



**Fig. 1.2** The number of awards per S&E baccalaureate has shifted downwards.

*Source:* NSF-DGE, Cumulative Index and related Datasets. Bachelor’s degree data tabulated by National Science Foundation/Division of Science Resources Studies; data from Department of Education.

Universities complained about the negative effects of a support gap between NSF fellows and other students. The 2003 COV recommended that the “NSF thoroughly examine the external consequences of the \$27,500 stipend and what intended and unintended consequences this might engender” (COV 2003, 13). For the 2005 fiscal year, the stipend has been set at \$30,000, with an accompanying \$10,500 cost of education allowance. The roller-coaster behavior of the award amount has prompted some in the graduate education community to seek a consistent basis for the determination of stipend levels.

The absolute number of GRF awards of about 900 to 1,000 in recent years is comparable to the levels in the 1950s. Relative to the increased supply of S&E bachelor’s graduates, however, the number of awards has fallen sharply. In 1952 to 1956 NSF awarded 5.4 GRFs per thousand S&E bachelor’s degrees. In 2000 to 2004 NSF awarded 2.2 GRFs per thousand S&E bachelor’s degrees—a 60 percent drop. Figure 1.2, which graphs the number of awards per S&E bachelor’s, shows that the drop occurred discontinuously around 1970.<sup>6</sup> It also shows sizable fluctuations in the ratio of

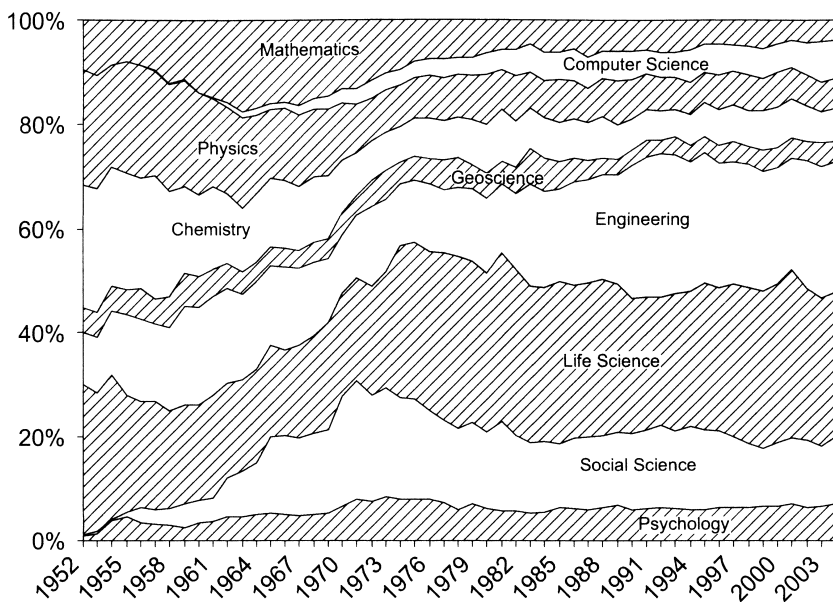
6. Since only about 7 percent of U.S. bachelor’s degrees in science and engineering go to the foreign-born, with little trend over time, it makes little difference if one uses all bachelor’s graduates or only nonforeign-born bachelor’s as the base for this calculation.

awardees to BS graduates due to changing numbers of awards granted before 1970.

### 1.1.1 Field and Demographic Distribution of Awards

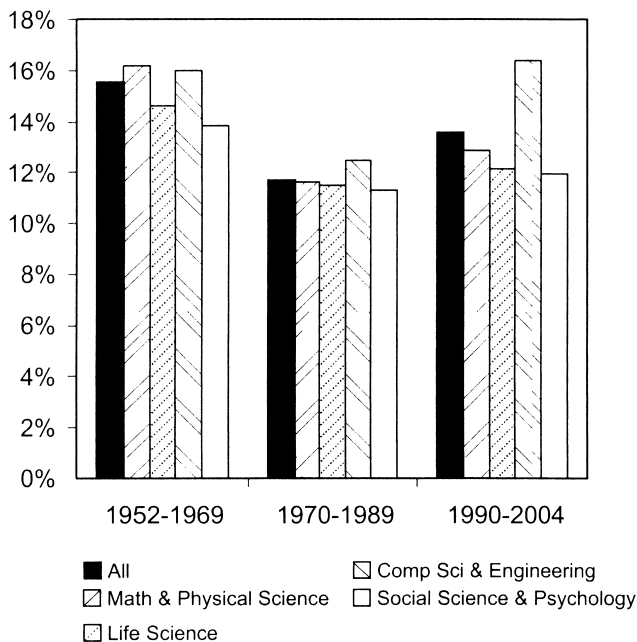
In the early years of the program, GRF awards went largely to physical sciences and mathematics and disproportionately to white men. Today, larger shares go to biological sciences, social sciences, and engineering, and also to women and minorities. Figure 1.3 shows the huge change in the distribution of awards among fields. From 1952 to 1956 the physical sciences obtained 49 percent of the awards, math received 9 percent, and life sciences obtained 25 percent compared to 4 percent in psychology and social sciences and 13 percent in engineering. In 2004, by contrast, 15 percent of awards were given to the physical sciences, 11 percent to mathematics and computer sciences, and 27 percent to life sciences; 30 percent were given in engineering and 17 percent in social sciences and psychology.

Given NSF policy to award approximately the same proportion of applicants among fields, the shift in awards among disciplines largely reflects a change in the distribution of applications among disciplines. Consistent with this, figure 1.4 shows that before 1990 the proportion of applicants who won awards in the major discipline categories were within a few percentage points of each other, although social science and psychology were



**Fig. 1.3** The GRF disciplinary distribution over time of applicants

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related datasets.



**Fig. 1.4 Percentage of applicants who gain GRF awards, selected years**

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related datasets.

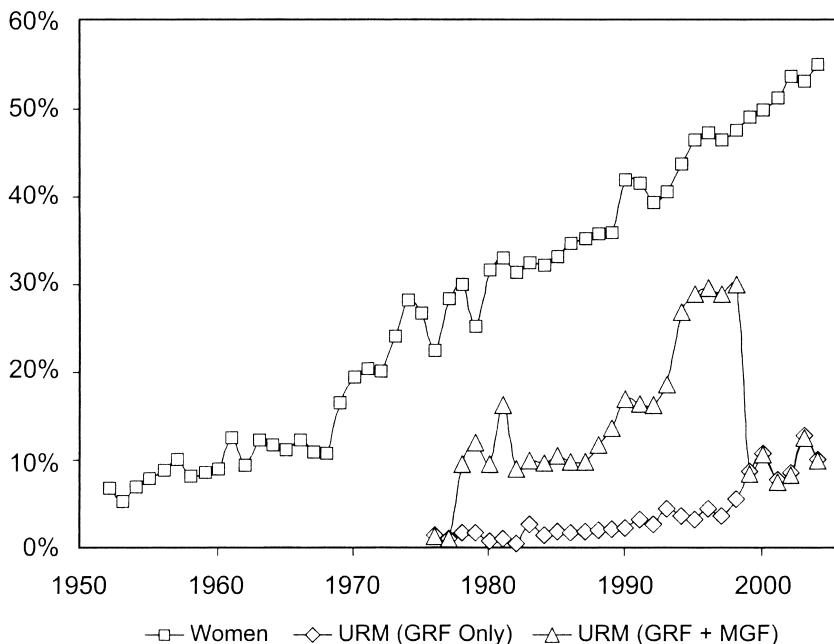
slightly disfavored. This is still the case after 1990, except that applicants in engineering and computer science have a higher chance of getting an award due to the process by which the NSF seeks to attract women to these disciplines.

Turning to the demographic characteristics of awardees, figure 1.5 shows a sharp rise in the proportion of women and underrepresented minorities<sup>7</sup> winning the GRF since the program was initiated. The female proportion of awardees climbed from about 5 percent in the 1950s to 55 percent of awards in 2004, with considerable variation among fields. Women make up roughly three-quarters of the awardees in life sciences, social science, or psychology compared to 20 percent of awardees in physical science or computer science. Around 1960, the female proportion in life science was about 25 percent, and in physical science it was around 6 percent, while there were barely any female engineering fellows.<sup>8</sup> To increase the female representation in these fields, NSF set up the Women in Engineering Program (WENG) in 1990, which became the Women in Engineering and

7. Underrepresented Minorities in S&E include African Americans, Hispanics, and Native Americans. Asian Americans are overrepresented in science and engineering.

8. In 1960 there were few computer science fellows, either male or female.





**Fig. 1.5 Percentages of fellows who are women or URMs**

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related datasets.

Computer Science (WECS) program in 1993. This program boosts the number of females by awarding fellowships to all or nearly all computer science or engineering women in NSF's quality group 2, and by making women in quality group 3 also eligible for awards if enough slots are available.<sup>9</sup> The data in figure 1.4 show that in 2001 applicants in computer science and engineering had a higher ratio of awards to applicants than other disciplines as a result of this program.

The data in figure 1.5 on the proportion of underrepresented minorities (URM) among GRF fellows starts in 1976 due to lack of ethnicity data before then. The GRF data is further complicated by the development in 1978 of the Minority Graduate Fellowship Program (MGF), which gave awards solely to underrepresented minorities, and its dissolution in 1998. The MGF was officially a separate program, although executed in parallel with the GRF. In 1999, the first program year since 1975 without the MGF, the NSF emphasized procedural changes in its review panel process in the

9. As described following, NSF categorizes applicants by group depending on their measured scholastic skills. All quality group 1 applicants are offered GRFs, and a sizable proportion of group 2 applicants are also given GRFs.

hope of maintaining a diverse awardee pool, partly by attracting more applicants through expanded outreach, and partly via other measures such as insuring a diverse group of panelists (COV 2003). Combining the fellows from the GRF and the MGF, the URM proportion is around 10 percent in the early 1980s, and climbed to 30 percent in 1998, the last year of the program. After that, the percentage dropped to the GRF-only proportion of around 10 percent, which is still much larger than the 1 to 2 percent two decades earlier. The disciplinary variation for URM fellows shows a disproportionately large proportion in social science and psychology.

The increased diversity in GRFs reflects two factors. First is that as a result of these programs and efforts, the likelihood that a woman or minority applicant would get an award has risen. In the first decade of the GRF, the overall award rate<sup>10</sup> was 17 percent while the award rate for women was 12 percent, as shown in figure 1.6. By the mid-1990s the female award rate exceeded the overall rate. The award rate for URMs in the GRF was low at 3 percent during the run of the MGF. The combined rate of URMs in both the GRF and MGF topped 18 percent from 1994 to 1998, which is actually greater than the MGF-only award rate, 15.3 percent, for the same time period.<sup>11</sup> After the end of that program, the number of URM awardees per applicants approached the overall rate of awardees per applicant, though it is still noticeably lower by several percentage points.

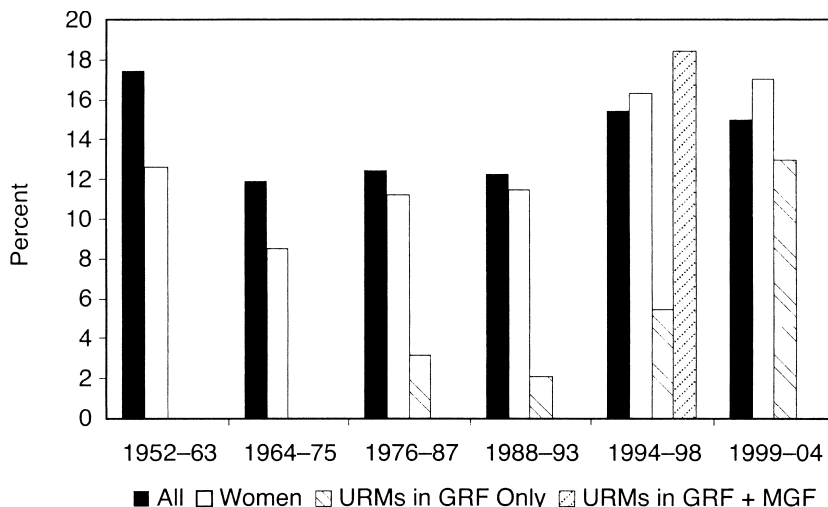
The second and more important reason for the increased diversity of GRFs is that an increasing share of bachelor's degrees in science and engineering have gone to women and minorities. This has raised the proportion of applicants from these groups, which in turn produces an increased share of awards, coupled with the NSF's efforts to improve diversity.

## 1.2 Determinants of GRF Awards

The GRF is a merit-based award. Each year, the NSF assembles a panel of experts to evaluate and rate applications and arrange them into quality groups. Quality Group 1 (QG1) is the highest quality group, and all applicants deemed fit for this group are offered the fellowship. Quality Group 2 (QG2) applicants are regarded as sufficiently meritorious for the fellowship, but not all in this category will receive the award. Applicants in other quality groups rarely receive the fellowship, but may receive an honorable mention, which carries some nonmonetary prestige. As with other aspects of the program, the procedures for determining awardees have evolved. Presently, panelists meet in groups arranged by discipline for an intensive two to three

10. “Award rate” is synonymous with “awardee to applicant ratio.”

11. The explanation for this result lies in the high incidence of duplicate applications. Data was available from 1994 to 1998 to match applicants in the two programs, and we found that approximately 90 percent of those who applied for the MGF also submitted an application for the GRF.



**Fig. 1.6** Average percentage of applicants winning awards in selected time periods, by broad demographic group

Source: NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related datasets.

Notes: Data on ethnic demography only available beginning in 1976. From 1994–1998, the available data allows the identification of duplicate applicants to both the MGF and GRF. To calculate “URMs in GRF+ MGF” the total number of unique fellows in either program was divided by the total number of unique applicants. The final year of the MGF program was 1998.

day session, which includes a briefing on the GRF program and a calibration exercise using applications from previous years. The categorization process consists of cycles during which the panelists score the applications on a scale from 1.00 to 4.99 (the lower number indicating higher quality), and then rank them. In the first cycle, at least two panelists score each application. Any applicants ranked below the sixty-fourth percentile are put in QG4 and no longer considered for the fellowship. A third panelist scores remaining applications and the applications are reranked. Approximately 55 percent of the awardees are from QG1, all of whom receive an award. Virtually all of the other awards are from QG2, with approximately 60 percent of QG2 applicants offered the fellowship (NSF 2002).

From 1998 to the present, NSF has instructed panelists to evaluate applications based on the two broad criteria common to NSF solicitations: *intellectual merit* and *broader impacts*. The panelists are given rating sheets upon which intellectual merit is broken down into subcriteria: GPA, GRE scores, references, proposed research, and previous research experience. Broader impacts are based on integration of research and education, diversity, benefit to society, and enhancement of scientific and technical un-

derstanding. The NSF emphasizes a comprehensive concept of merit, and advises panelists to avoid placing too much emphasis on the easily quantifiable quality measures such as GRE scores and GPA. In the latest program years, GRE scores are not required on the GRF application. Further, there is a new emphasis on the value of leadership skills in NSF fellows (COV 2003). It is possible that future applicants will be asked to provide evidence of leadership potential, and future panelists will include this in their assessment of merit.

Because the NSF recognizes the value of a diverse workforce, the program implements measures to insure representation in geography, demographics, and discipline. Specifically, the applicants from Quality Group 2 who receive the fellowship are chosen based on an algorithm that includes their rank within QG2, but also other factors. The key policy with respect to discipline is that NSF allocates the number of QG1 and QG2 positions to review panels on the basis of the disciplinary distribution of all applicants and the number of awards NSF plans to offer.

### 1.2.1 Regression Analysis

To see how the selection procedure uses the information on student skills to determine which applicants win awards, we estimated linear probability models in which the dependent variable is the 0/1 measure of obtaining an award and the explanatory variables consist of measures of student skills from the NSF’s Cumulative Index data set. The CI has recorded the scores that a panelist gives to an application starting in 1976, but does not record assessments of the subcriteria, which limits our choice of measures largely to the quantitative and verbal GRE scores, the GPA, and the scoring of reference letters. Table 1.1 lists the measures of scholastic skills in the data and gives their mean values and standard deviations, as well as the overall mean of the probability of winning the GRF award, and the scales used to measure each quantity. The GRE scores are coded on a 200 to 800 scale; the GPA uses a four point scale; reference scores were originally scaled from ten to seventy, with lower scores being better; finally the Panel Ratings<sup>12</sup> (which reflect the GRE, GPA, and reference scores) were originally coded on a scale of 100 to 600, again with lower score representing higher quality applicants. The panel ratings and reference scores changed scales at various times in GRF history. To deal with this, we normalized the panel and reference data such that they have zero mean and standard deviation of one in the applicant pool of each year. We also reverse-coded the actual panel and reference scores, so that increases in all variables imply a more favorable situation. Regression coefficients for GRE scores are sometimes re-

12. In the original Cumulative Index data set, each applicant may have more than one panel rating if more than one panelist reviewed the application. An *average panel rating* was trivially generated for each applicant. In this chapter, “panel rating” means “average panel rating.” Similarly, “reference score” is synonymous with “average reference score.”

**Table 1.1** Means, standard deviations, and scaling of the quantitative measures of applicant scholastic skills

Summary for key data in the GRF datafiles	Total time period	Records	Overall award rate	
	1952–2004	271,391	13.50%	
Data available	Mean	STD	Full range	
Panel ratings	1976–1993	328	115	100 to 600
	1994–1998	315	118	100 to 600
	1999	3.11	1.02	1.00 to 6.00
	2000–2004	2.90	0.85	1.00 to 4.99
Reference score	1954–1993	24.2	8.7	10 to 70
	1994–1998	1.91	0.65	1 to 7
GRE quantitative	1952–2004	691	100	200 to 800
GRE verbal	1952–2004	606	104	200 to 800
GPA	1955–1998	3.49	0.40	1 to 4
Ethnicity	1976–2004	(1, 2, 3, 4, 5)		
	1981–1987	(1, 2, 3, 4, 5, 6, 7, 8, 9)		
	1988–2000	(A, B, C, D, E, F, H, I, J)		
	2000–2004	(t, f) or (y, n, b)		
Gender	1952–2000	(m, f)		
	2001–2004	(m, f, b)		

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets.

*Notes:* For analysis, Panel ratings and Reference scores will be normalized such that they have mean zero and standard deviation one in the applicant pool of each year. These variables are also recoded such that higher values denote better ratings. For GRE quantitative and GRE verbal, values in original data set above 800 set to 800. For GPA, full range (0 to 3) before 1971, values shifted by + 1. A separate variable codes each ethnic category. For Gender, value of “b” indicates “decline to answer.”

ported in tables with the scores scaled as GREQuant/100 or GREVerbal/100, which effectively puts them on a two to eight point scale.<sup>13</sup> Note in the table that applicants have higher GRE quantitative scores than they have GRE verbal scores, so that while the standard deviations of scores are similar, the coefficient of variation is lower for the quantitative score than for the verbal score.

Table 1.2 records linear regressions of the impact of scholastic attributes, demographic factors, and field dummies for nearly the full data set from 1955 to 1998, which lacks measures of minority status, panel scores, and reference letters; and for the 1976 to 1998 period when we have data on those measures as well as the others. The column (1) and (2) regressions do not include the panel rating, while the column (3) regression includes the panel rating. The column (1) estimates show that over the entire period, GRE Quantitative, GRE Verbal, and GPAs have sizable and significant

13. This allows us to report coefficients concisely as 0.XX rather than 0.00XX.

**Table 1.2** Linear probability model for the impact of scholastic and demographic variables on GRF awards

	Offered Award		
	1955–1998	1976–1998	1976–1998
Panel rating	—	—	0.170 (0.001)
Reference score	0.082 (0.001)	0.074 (0.001)	0.022 (0.001)
GRE quantitative/100	0.034 (0.001)	0.035 (0.001)	–0.010 (0.001)
GRE verbal/100	0.060 (0.001)	0.066 (0.001)	0.019 (0.001)
GPA	0.089 (0.002)	0.095 (0.003)	–0.010 (0.003)
Female	—	0.033 (0.002)	0.037 (0.002)
Minority	—	0.087 (0.004)	0.077 (0.003)
Field effects:			
Chemistry	–0.003 (0.003)	0.011 (0.004)	0.045 (0.004)
Computer science	–0.007 (0.004)	0.009 (0.005)	0.115 (0.005)
Engineering	0.022 (0.003)	0.048 (0.004)	0.093 (0.003)
Earth/atmospheric	0.033 (0.004)	0.034 (0.005)	0.049 (0.005)
Life science	0.037 (0.002)	0.034 (0.003)	0.040 (0.003)
Math	–0.019 (0.003)	–0.012 (0.005)	0.051 (0.005)
Physics/astronomy	–0.030 (0.003)	–0.006 (0.005)	0.060 (0.004)
Psychology	0.000 (0.003)	0.005 (0.005)	0.018 (0.004)
Social science	—	—	—
Year effects	Yes	Yes	Yes
Observations	207,498	107,658	107,597
R <sup>2</sup>	0.198	0.1793	0.2731

Source: NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets.

effects on the probability of obtaining the GRF. A 100 point increase in Verbal GRE score raises the probability of award by 6.4 percent. This might not seem large, but since just 13.5 percent of applicants get an award this raises the probability of getting an award by nearly 50 percent. The higher effect of the Verbal GRE than of the Quantitative GRE may seem

surprising, given the presumably greater importance of quantitative thinking in science and engineering. It reflects the fact that the Verbal GRE is often the key distinguishing factor between applicants with similarly high Quantitative GREs. Finally, the calculations show that women and minorities have higher chances of getting an award than other applicants. The field dummy variables show some variation in the probability of getting an award among fields. Engineering, life science, and earth/atmospheric science have significantly higher probabilities of getting an award while mathematics has a lower one compared to the omitted social science group.

The results in column (2), which include the measure of the quality of references and dummy variables for gender and minority status for the period 1976 to 1998 are similar: GREs, the GPA, and the reference scores are key determinants of the probability of getting an award, while women and minorities have a higher probability of gaining awards. The dummy variables for fields show moderate disciplinary variation. Column (3) adds the panel rating to the list of variables. Since the NSF panels weigh heavily the effect of GREs, GPAs, and reference scores in deciding which applicants should be given awards, inclusion of the panel rating alters the estimated coefficients markedly. The coefficients on factors on which panels place most weight in their evaluation fall relative to those that enter less highly in the panel rating. The GRE quantitative score and the GPA obtain negative coefficients conditional on the panel rating, implying that the panels weigh those factors especially heavily in their ratings. In addition, conditional on the panel ratings the field dummies become large and significant relative to the omitted social science group. This presumably reflects the initial allocation of the number of QG1 and QG2 positions on the basis of the disciplinary distribution of applicants, which leads panels in the quantitative and physical sciences to give lower ratings than those in the social sciences to applicants with comparable measured scholastic skills. From this perspective, the estimated positive coefficients on the field dummies in table 1.1 are a correction that takes account of the attributes of applicants in these fields.

Consistent with this interpretation, table 1.3 shows that panel ratings are indeed lower for applicants with the same measured scholastic skills in physical sciences than in the social sciences. For instance, the regression shows that an applicant in chemistry obtains a  $-0.198$  lower panel rating than an applicant in the omitted social science group *with the same GREs, GPA, reference score, and demographics*. The rank correlation of the ordering of fields between the table 1.3 estimates of the panel ratings and the ordering of fields in the column (3) of table 1.2 is  $-0.92$  (inclusive of the omitted group). Thus, the modest estimated impacts of fields in columns (1) and (2) of table 1.2 reflect two offsetting factors: lower panel ratings in physical sciences but an offsetting higher probability of getting an award conditional on panel ratings in those disciplines.

**Table 1.3** Regression coefficients and standard errors for scholastic and demographic determinants of average panel rating, 1976–1998

Average panel rating 1976–1998	
Reference score	0.304 (0.002)
GRE quantitative/100	0.268 (0.003)
GRE verbal/100	0.279 (0.002)
GPA	0.614 (0.006)
Female	−0.024 (0.004)
Minority	0.057 (0.007)
Field effects:	
Chemistry	−0.198 (0.009)
Computer science	−0.629 (0.010)
Engineering	−0.263 (0.007)
Earth/atmospheric	−0.092 (0.011)
Life science	−0.038 (0.006)
Math	−0.367 (0.010)
Physics/astronomy	−0.386 (0.009)
Psychology	−0.079 (0.009)
Social science	—
Year effects	Yes
Observations	107,597
$R^2$	0.6079

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets.

### 1.2.2 The Effect of Demographics

As noted, the table 1.2 regressions for obtaining a GRF show that women and minorities have a greater probability of getting an award than do majority men with similar measured scholastic skills. This is consistent with the goal of increasing the diversity of the S&E workforce. The table 1.3 regression of the panel ratings on scholastic skills and demographic characteristics show, however, that panels give lower scores to women than to majority men with the same measured attributes while they give higher



scores to minorities than to majority men with the same measured attributes. Since the probability of getting an award is higher for both women and minorities than for majority men, the difference in coefficients in panel ratings suggests that the process of taking account of diversity differs between the two groups.

To unpack this difference, we estimated a recursive model. First, we related the panel scores and the probability of being assigned to groups 1 and 2 to demographic characteristics. Then we examined the probability of getting an award conditional on being in groups 1 and 2 and on the panel rating. Table 1.4 gives the results of this analysis for 1990 to 1997 and 1998 to 2004. The data for 1990 to 1997 includes measures of GRE scores, GPA,

**Table 1.4** Regression coefficients for determinants of panel rating, quality group, and award offer, 1990–2004

	1990–1997			1998–2004		
	Panel rating (1)	Assigned quality group 1 or 2 (2)	Award conditional on being group 1 or 2 (3)	Panel rating (4)	Assigned quality group 1 or 2 (5)	Award conditional on being group 1 or 2 (6)
Female	–0.012 (0.006)	–0.021 (0.003)	0.38 (0.014)	0.135 (0.009)	0.024 (0.004)	0.25 (0.012)
Minority	–0.003 (0.009)	0.076 (0.004)	0.072 (0.041)	0.161 (0.015)	0.091 (0.006)	0.073 (0.021)
Panel rating	—	—	0.275 (0.039)	—	—	0.546 (0.028)
Quantitative/100	0.263 (0.004)	0.035 (0.002)	—	0.334 (0.007)	0.081 (0.003)	—
Verbal/ 100	0.235 (0.003)	0.074 (0.002)	—	0.285 (0.005)	0.093 (0.002)	—
GPA	0.634 (0.010)	0.118 (0.004)	—	—	—	—
Reference	0.289 (0.003)	0.093 (0.002)	—	—	—	—
Field effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	47,821	47,851	4,904	41,404	41,404	5,695
R <sup>2</sup>	0.5882	0.2109	0.1445	0.215	0.0901	0.1574

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets.

*Notes:* For comparison with the 1998–2004 analyses, we estimated the 1990–1997 equations excluding GPA and reference scores. The coefficients (standard error) on the dummy for female barely changed: it was –0.012 (.006) in the panel rating equation and –0.026 (.003) in the equation for being assigned to group 1 or 2. For minorities, the coefficient on the dummy for minority status in the panel rating became –0.241 (.011). The implication is that GPA and reference scores for minority applicants were much lower than for others. But the probability of being assigned to group 1 or 2 gave to minorities a positive, though smaller, coefficient than in the table, .018(.003).

and a measure of the quality of reference letters. The 1998 to 2004 data do not include GPA or the quality of reference letters, and cover the period when NSF revised its merit criteria. Column (1) shows that in 1990 to 1997 women received a slightly lower score than men from review panels while minorities received comparable panel ratings as the majority group. Women also were less likely to be assigned to quality groups 1 and 2 than men, while by contrast minority applicants were more likely to be assigned to those groups. Among persons in groups 1 and 2, and conditional on panel ratings, women were much more likely to obtain awards than men while minority applicants had a modestly higher probability of obtaining awards. Thus, the higher probability of women getting awards occurs at the last stage, presumably in part because the WECS (WENG) program boosted the chances that women in quality group 2 would get GRFs, whereas the higher probability of minorities getting awards occurs because panels were more likely to assign them into groups 1 or 2. To make sure that this interpretation is not marred by the availability of GPA and reference scores in the 1998 to 2004 period but not earlier, we estimated the 1990 to 1997 equations excluding those variables. As noted in the table note, the coefficient on females barely changed with their exclusion.

The 1998 to 2004 regressions show a very different pattern, indicating that the change in NSF’s policies in awarding GRFs had a substantial effect on the awards process. Column (4) shows that panels give both women and minorities higher ratings in 1998 to 2004 than in 1990 to 1997. Column (5) shows that panels are more likely to assign them to quality groups 1 or 2 than in 1990 to 1997. The coefficient on being female in column (6) shows that, conditional on being in groups 1 or 2, the boost given to women in the probability of getting an award was more moderate in 1998 to 2004 than in 1990 to 1997, presumably because they were getting higher panel ratings and having a higher chance of being assigned to quality groups 1 or 2. By contrast, the coefficient on being minority in column (6) is nearly identical to the coefficient in column (3). Finally, panel ratings have a stronger impact on awards in 1998 to 2004 than in 1990 to 1997. In sum, table 1.4 provides evidence for the change in NSF’s emphasis that was enacted, at least partly, to increase diversity in S&E fields.

Finally, table 1.5 uses a regression model to examine the GRE scores of applicants with different demographic characteristics from 1976 to 2004—a period in which the proportion of applicants who were women or minority increased substantially. In 2004 there were 8,939 applicants, of whom 46 percent were women and 10 percent underrepresented minorities. In 1976, there were 5,366 applicants, of whom 30 percent were women and 4 percent were underrepresented minorities. To the extent that increased numbers of applicants draw persons with decreasing measured scholastic skills, one would expect that the increased proportion of women and minorities would reduce the average scores of those groups.

The coefficients and standard errors on the demographic variables in the

**Table 1.5 Demographic correlates of test scores, regression coefficients and standard errors, 1976–2004**

	GRE Quantitative test scores					GRE Verbal test scores						
	1976–1983	1984–1991	1992–1998	1999–2004	1976–1983	1984–1991	1992–1998	1999–2004	1976–1983	1984–1991	1992–1998	1999–2004
Female	-41.7 (1.0)	-36.3 (0.9)	-35.8 (0.8)	-28.9 (0.8)	0.0 (1.2)	-10.6 (1.0)	-16.1 (1.0)	-7.2 (1.0)				
Black	-175.9 (3.2)	-166.5 (1.8)	-146.3 (1.4)	-100.1 (1.9)	-168.7 (3.8)	-160.1 (2.1)	-146.5 (1.7)	-100.4 (2.4)				
Hispanic	-77.5 (3.5)	-98.0 (2.0)	-87.1 (1.5)	-54.7 (1.7)	-80.5 (4.2)	-107.8 (2.3)	-95.9 (1.9)	-70.6 (2.2)				
Other minority	-49.2 (8.1)	-73.9 (4.1)	-69.1 (3.1)	-44.7 (3.4)	-41.0 (9.5)	-69.8 (4.9)	-60.7 (3.8)	-38.7 (4.4)				
Field effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes				
Observations	30,213	41,045	41,804	36,658	30,213	41,045	41,804	36,659				
$R^2$	0.3256	0.3688	0.3778	0.2536	0.1116	0.1868	0.2273	0.1279				

Source: NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets.

table show the opposite pattern, particularly for GRE quantitative scores. The scores of both women and minority applicants rise relative to those of majority men, though both groups still score lower. The coefficient on the dummy for women rises from  $-41.7$  in 1976 to 1983 to  $-28.9$  in 1999 to 2004. The coefficient on being black rises from  $-175.9$  in 1976 to 1984 to  $-100.1$  in 1999 to 2004. The negative coefficient on being Hispanic changes less evenly but is lower in 1999 to 2004 than in 1976 to 1983, while that for other minorities shows little change. The regression coefficients for GRE verbal scores tell a similar story, with the exception that women applicants go from having the same verbal scores as men in 1976 to 1983 to a modest 7.2 lower score in 1999 to 2004. The coefficient for being black, however, advances steadily from  $-168.7$  to  $-100.4$ . Overall, the results show that the increased proportion of female and minority GRFs was associated with generally improved measured skills. The implication is that the GRF program is attracting increasingly qualified women and minorities.

### 1.3 Supply Responses of GRF Applicants

The NSF has two policy levers that can affect the supply of applicants to the GRF program and, given the process of granting awards, that can determine the measured skills of awardees: the dollar amount of the awards and the number of awards.

*Increasing the dollar value of awards* should increase the number of applicants. It could increase or decrease the average measured skill of *applicants* depending on whether they come largely from persons with the highest skills or from those with lower skills. But given the process of selecting awardees, increases in the value of awards should increase the average skill of *awardees* as long as the increased value of awards attract some persons with high skills. This is because review panels would be expected to select high skill candidates from the pool of newly attracted persons over less skilled candidates that they would have selected from the smaller initial pool of applicants.

*Increasing the number of awards*, by contrast, is likely to reduce the measured quality of awardees since it means accepting applicants who would otherwise be rejected. The key issue here is the extent to which the increase in numbers reduces quality, which depends on the number of applicants on the margin of acceptance with similar skills. If there is a large number of persons on the margin with similar skills, then increasing the numbers of awards will reduce measured skills modestly. If there is a small number, an increased number of awards could reduce measured skills substantially.

#### 1.3.1 Number of Applicants

To examine the link between the value of stipends and the pool of GRF applicants, we calculated a *relative value of stipends* by taking the ratio of the

dollar value of GRF awards to the average annual earnings of college graduates aged twenty-one to twenty-five<sup>14</sup> and contrasted this with the *relative number of applicants* by dividing the number of applicants by the number of S&E bachelor graduates. As figure 1.7 shows, the two series track each other closely. For example, the sharp increase in the stipend value in the 1980s, when NSF increased stipend amounts from \$3,900 (1979) to \$14,000 (1991) is matched by a large rise in the number of applicants relative to S&E bachelor's graduates. By contrast, from 1991 to 1999 the relative value of stipends fell, as the nominal stipend amount rose by \$1,000 while college graduate earnings rose by about \$7,600; and the relative number of applicants fell. Since 1999 the nominal amount of the stipend has doubled, and the number of applications has nearly doubled as well.<sup>15</sup>

Another factor that might affect the decision to seek a GRF and go on into graduate school is the ease of finding work with a bachelor's degree, which varies with the business cycle. Figure 1.8 compares the relative number of applicants to the unemployment rate of college graduates.<sup>16</sup> Since the late 1980s, the size of the GRF applicant pool has moved countercyclically: a weak labor market for college graduates generated a higher number of applicants as a share of bachelor's degrees. But in the early to mid-1980s, the unemployment rate fell while the number of applicants rose, presumably in part because of the increase in stipend amounts shown in figure 1.7. Thus, figure 1.8 suggests that unavailability of jobs for recent college graduates is a secondary factor influencing the decision to apply for a GRF stipend.

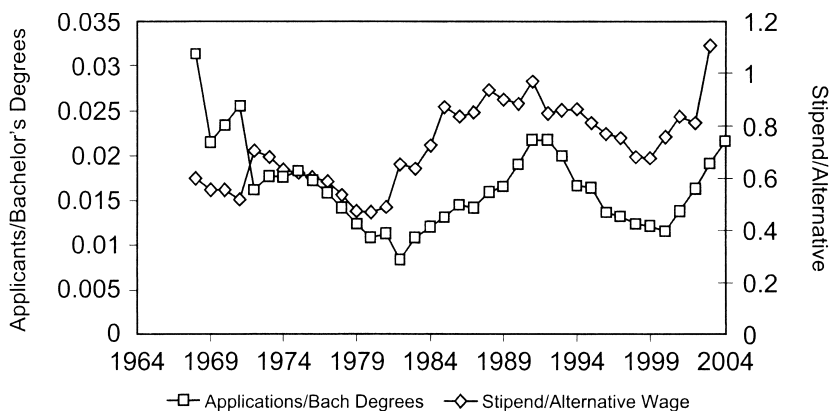
Taking the analysis a step further, we amalgamated the data on GRF fellowships into nine major fields<sup>17</sup> by year from 1952 to 2004, to obtain a cross-field time series panel with 477 observations. Using this sample, we did a multivariate regression analysis of the determinants of the number of GRF applicants in each field and year. Since the NSF changes the stipend value in particular years by policy decision rather than changing it in response to the previous years' labor market, we take these changes as exogenous and use least squares to estimate the effect of relative stipend values on the relative number of applicants. To control for field differences and changes in market conditions by field over time, the regressions in-

14. We estimated the college graduate earnings from the Integrated Public Use Microdata Series (PUMS) of the March Current Population Survey. College graduates include those who obtain degrees higher than a bachelor's degree, but the age range is restrictive enough to exclude most doctorate recipients.

15. Bachelor's degree data for 2002 to 2004 are extrapolated because actual figures are not available. The key to the change in the ratio of applicants to bachelor's graduates is that the number of applicants jumped from 4,852 in 1998 to 8,939 in 2004.

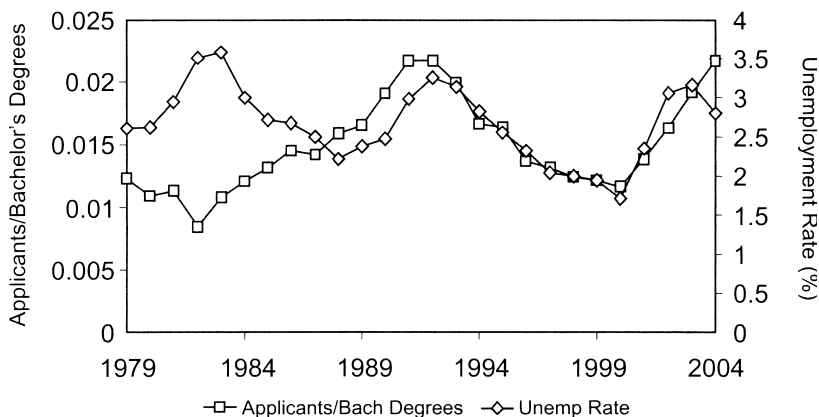
16. Estimated from the annual Current Population Survey (CPS) Outgoing Rotation Group.

17. The nine fields are: Chemistry, Computer Science, Earth/Atmospheric Science, Engineering, Life Science, Mathematics Physics, Psychology, and Social Science. The availability of earnings data for college graduates and degree data by field limits the sample to 1968 to 2001.



**Fig. 1.7 GRF applications relative to S&E bachelor's degrees and relative value of GRF stipend, 1968–2004**

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related datasets. Bachelor's degree data tabulated by National Science Foundation/Division of Science Resources Studies; data from Department of Education.



**Fig. 1.8 GRF applications relative to S&E bachelor's degrees and college graduate unemployment rate, 1979–2004**

*Source:* NSF DGE, Cumulative Index of the GRF Program and related datasets. Bachelor's degree data tabulated by National Science Foundation/Division of Science Resources Studies; data from Department of Education. Unemployment data are estimated from the annual Current Population Survey (CPS) Outgoing Rotation Group, Bureau of Labor Statistics.

clude field dummy variables and separate time trends for the nine fields. In addition, the regressions include the number of bachelor's degrees earned in the nine fields in the previous year.<sup>18</sup>

18. We recognize that students may apply for a GRF in a different field than the field in which they earned their undergraduate degree, but still regard the number of undergraduate degrees as a reasonable indicator of interest in the field.

**Table 1.6** Determinants of the number of GRFP applicants, 1969–2004

	Dependent variable: ln(applicants in academic field in current year)			
	(1)	(2)	(3)	(4)
Log(bachelor's degrees) by field in current year	0.195 (0.057)	0.304 (0.063)	0.298 (0.062)	0.516 (0.066)
Log(stipend/outside salary) in previous year	0.996 (0.084)	0.916 (0.060)	0.852 (0.059)	0.772 (0.056)
Unemp. rate for college grads age 21–25	—	0.049 (0.013)	—	—
Unemp. rate for all college grads	—	—	0.104 (0.024)	0.094 (0.022)
Log(awards/bachelor's degrees) by field in previous year	—	—	—	0.349 (0.054)
Field effects	Yes	Yes	Yes	Yes
Field × time trend	Yes	Yes	Yes	Yes
Observations	324	234	234	234
R <sup>2</sup>	0.8931	0.955	0.9561	0.9634

*Source:* NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets, as described in text.

*Notes:* Outside salary are earnings of college graduates aged 21–25. Outside salaries and unemployment rates from Current Population Survey.

Table 1.6 displays the regression results. The basic specification in column (1) shows that a larger number of bachelor graduates in a field generates a larger number of applicants from that field; and that stipends in the previous year have a strongly positive effect on applications, with a near unitary elasticity. Column (2) adds to the set of explanatory variables the unemployment rate among college graduates twenty-one to twenty-five years of age as an indicator of the availability of alternative jobs for those contemplating going directly to the job market. The sample size shrinks because the unemployment variable is only available starting in 1979. Consistent with figure 1.8, the unemployment rate has a positive effect on the number of applications in column (2). In column (3) we replace the unemployment rate of twenty-one to twenty-five-year-olds with the unemployment rate for all college graduates regardless of age. The effect of unemployment nearly doubles: a one percentage point rise in the unemployment rate raises the GRFP applications by just over 10 percent, presumably because a one point increase in unemployment for all graduates is a bigger economic shock than a one point increase in unemployment for all graduates is a bigger economic shock than a one point increase for younger grad-

uates (whose unemployment rate will be more volatile). Finally, in column (4) we include a measure of the relative availability of awards by field—the number of awards divided by the number of bachelor’s graduates in the field lagged one period. This shows that more students apply when the probability of obtaining an award is perceived to be higher. All of the specifications show that the coefficient on the log bachelor’s degrees is substantially below unity. One likely reason for this is that the expansion of degrees in the United States has occurred largely at institutions of lower academic quality, so that the marginal bachelor’s graduate is unlikely to be of GRF quality. In sum, the regression specifications in table 1.6 show that the number of GRF applications responds substantially to the relative value of the stipend and to the perceived availability of awards. For every 10 percent increase in the stipend value, the number of applications goes up by 8 to 10 percent.

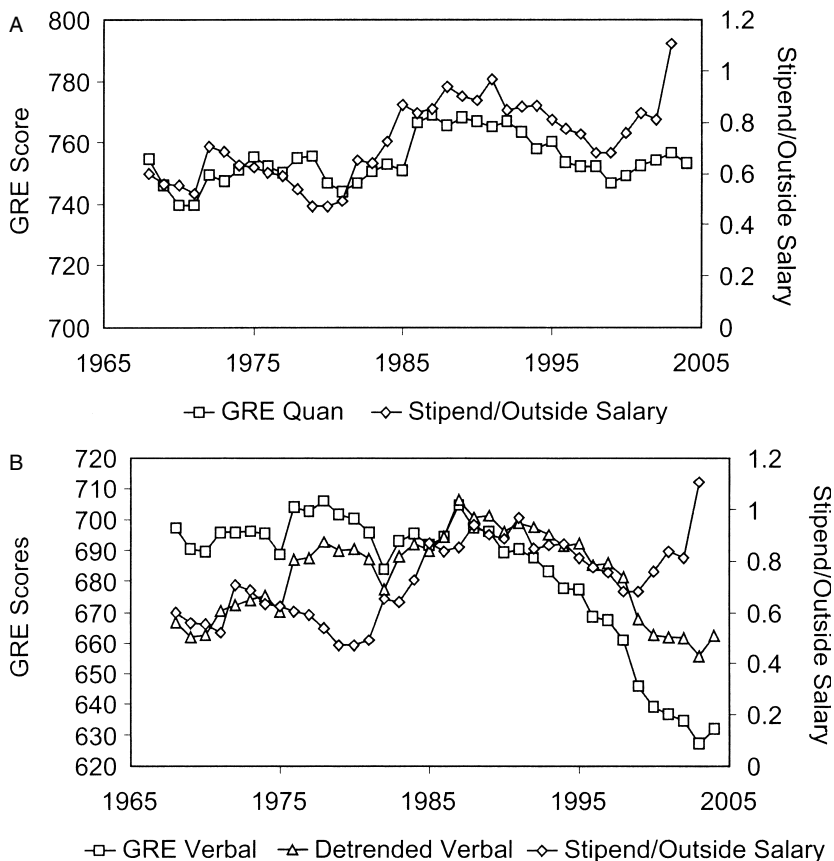
### 1.3.2 Measured Skills of Awardees

To see how the value of stipends affects the measured achievement of GRF awardees, we compare the GRE scores and GPAs of awardees over time and across fields. Figure 1.9 provides a first indication that stipend values affect the qualifications of the awardees. Figure 1.9, panel A shows that the average GRE quantitative score of awardees has been higher when the relative stipend value is higher. Figure 1.9, panel B shows a similar relation between stipend values and average GRE verbal scores of awardees. The time series of raw verbal scores trended downward beginning in 1987, so we examined the time series of GRE verbal scores adjusted for trend, as well as the raw time series. The trend-adjusted series shows a rise in scores with the large increases in relative stipend values in the first half of the decade, and a subsequent fall in verbal scores throughout the 1990s. However, the continuing decline in GRE verbal scores through the late 1990s/early 2000s cannot be explained by the evolution of stipend values, which are rapidly rising during this time.

To quantify the pictures in the figures, we regressed measures of the scholastic qualifications of awardees on the number and value of GRF awards, using a cross-field time series regression framework similar to that in table 1.6, with field fixed effects and field-specific time trends included in all regressions. Table 1.7 gives the results of this analysis.<sup>19</sup> The first row of coefficients shows that the number of awardees has a significant but quantitatively modest negative effect on the measures of quality. An approximate 10 percent or 0.10 ln increase in the number of awards is associated with a decline in average quantitative GRE scores by 1.1 points and a decline in verbal GRE scores by 2.2 points on the 200 to 800 GRE scale, and a drop

19. Freeman (2005) finds similar results with a different specification that suggests that the basic findings in tables 1.6 and 1.7 are robust to some changes in specifications.





**Fig. 1.9 Mean GRE scores of awardees and relative stipend value, 1968–2004: A, quantitative scores; B, verbal scores**

*Source:* NSF DGE, Cumulative Index of the GRF Program and related datasets. Salary data estimated from the Integrated Public Use Microdata Series (IPUMS) of the March Current Population Survey.

in average GPA by 0.009 points (on a one to four scale). These findings are consistent with the notion that as the number of fellowships increases, the review panels move down the measured scholastic skill distribution to give stipends to marginal awardees that are less outstanding than the inframarginal winners. The larger drop in the verbal scores than in the quantitative scores presumably reflects the wider variation in those scores and the extent to which panels view high quantitative scores as more critical to success in the sciences. The second row of coefficients in table 1.7 shows that, with the number of awards fixed, increases in the relative value of fellowships increases two of our measures of the skill of awardees, but has no discernible

**Table 1.7** Determinants of awardee achievement, 1969–2004

	GRE quantitative	GRE verbal	GPA
Log(number of awards) by field in current year	−10.6 (1.80)	−21.8 (3.30)	−0.087 (0.01)
Log(stipend/outside salary) in previous year	29 3.3	35.4 6.1	−0.007 0.017
Field effects	Yes	Yes	Yes
Field × time trend	Yes	Yes	Yes
Observations	324	324	270
R <sup>2</sup>	0.8943	0.684	0.7354

*Source:* Tabulated from NSF, Division of Graduate Education, Cumulative Index of the GRF Program and related data sets, as described in text.

*Note:* Outside salary are earnings of college graduates aged 21–25, tabulated from Current Population Survey.

effect on the third measure. An increase in the relative value of the stipend by 10 percent raises the average quantitative GRE score by 2.9 points and raises the average verbal GRE score by 3.5 points. However, the average GPA of awardees appears invariant to the stipend values.

Overall, the table 1.7 results on the GRE scores suggest that at least some of the students on the margin of applying for a GRF fellowship are high-achieving students who can be drawn away from other career or academic opportunities through the prospect of larger awards. In the context of the occupational choice model of Roy (1951), these results imply a positive correlation between ability in science and engineering and ability in alternative careers: the able S&E students who are drawn into the GRF program by higher stipends perform well on general measures of achievement (i.e., GRE scores) that are associated with success in other occupations.

To test this interpretation, we regressed the number of applicants with relatively high measured skills (defined as those scoring 770 or over on the GRE quantitative exam, 680 on the verbal GRE, and those with GPAs above 3.88 on a four point scale) on the relative value of stipends and the number of awards given. By controlling for the number of awards, we get a measure of the *relative* number of applicants with high skills, which should largely determine the measured skills of awardees. Column (1) of table 1.8 shows that a 10 percent increase in the stipend value raises the number of applicants with 770 or over on the quantitative GRE by 12.2 percent. Column (2) shows a comparable increase in the number of applicants with 680 or over on the verbal GRE. The column (3) regression gives a 0.448 impact of stipend values on the number of students with very high GPAs.<sup>20</sup>

20. We also estimated models with the log of the number of applicants in the regression. This gave smaller positive coefficients to the relative stipend value in the regressions for the

**Table 1.8** Determinants of the number of applicants scoring above high thresholds

	Log(applicants with GRE quantitative > 770) (1)	Log(applicants with GRE verbal > 680) (2)	Log(applicants with GPA > 3.88) (3)
Log(stipend/outside salary) in previous year	1.221 (0.104)	1.276 (0.086)	0.448 (0.100)
Log(number of awards) by field in current year	0.539 (0.090)	0.664 (0.074)	0.512 (0.077)
Field effects	Yes	Yes	Yes
Field $\times$ time trend	Yes	Yes	Yes
Observations	261	261	207
$R^2$	0.9101	0.9097	0.9235

*Source:* Tabulated from NSF, Division of Graduate Education, Cumulative Index of the GRF program and related data sets, as described in text.

*Note:* Outside salary are earnings of college graduates aged 21–25, tabulated from Current Population Survey.

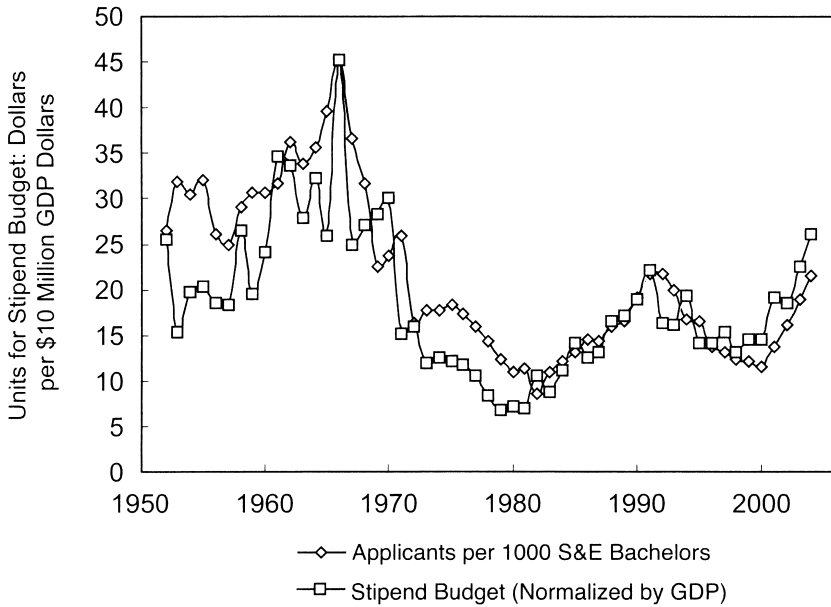
In sum, higher stipend values incentivize enough high-achieving students to seek GRF offers to enable the NSF selection procedure to increase the measured skills of awardees when stipend values are higher.

### 1.3.3 The Economic Supply Link

Overall, our results show that the resources NSF allocates to stipends, through the number of awards and their value, affects the supply of applicants, both in terms of numbers and measured scholastic skills. As a summary of the economic supply link, we have graphed in figure 1.10 our measure of the supply of students seeking GRFs—the number of applicants relative to the number of S&E baccalaureates—and a measure of NSF’s stipend budget (the number of awards offered in a particular year multiplied by the stipend amount) relative to GDP, which we take as the broadest possible measure of other economic activity. The stipend budget is an indicator of national resources spent to make graduate S&E study more financially attractive, whereas GDP represents national resources broadly. The tight link between the two series gives the bottom line message: if the United States wants more applicants for S&E fellowships and thus potentially more graduate students and Ph.D.s in these fields, the country has to spend more money on this objective.

---

number of applicants with high math and verbal GREs, because the number of applicants is dependent on stipend values and associated with the number of applicants with high scores. It also raised the number with high GPAs, but conditional on the number of applicants, the relative value of stipends were modestly negatively associated with the number of applicants with high GPAs.



**Fig. 1.10 Fraction of bachelors choosing to apply to GRF vs. total GRF stipend budget/GDP**

*Source:* NSF DGE, Cumulative Index of the GRF Program and related datasets. Data on the Gross Domestic Product (GDP) from the Bureau of Economic Analysis, an agency of the U.S. Department of Commerce.

## 1.4 Alternative Policy Scenarios

The analyses thus far have shown that the two policy levers—numbers of GRFs and the dollar value of the awards—affect the number and measured skills of persons seeking GRF awards and that coupled with NSF selection procedures, they determine the measured skills of awardees. In this section, we consider three scenarios that assess the way changes in the mode of determining stipend policy could affect the supply of students seeking and obtaining GRFs.

1. Agency determined changes in the number and value of awards.
2. Granting awards and setting values through fixed rules rather than policy discretion.
3. Allocating the number of quality group 1 and 2 positions among disciplines on the basis of quantifiable academic skills of applicants rather than on the number of applicants.

### 1.4.1 Agency Determined Changes in Stipend Numbers and Values

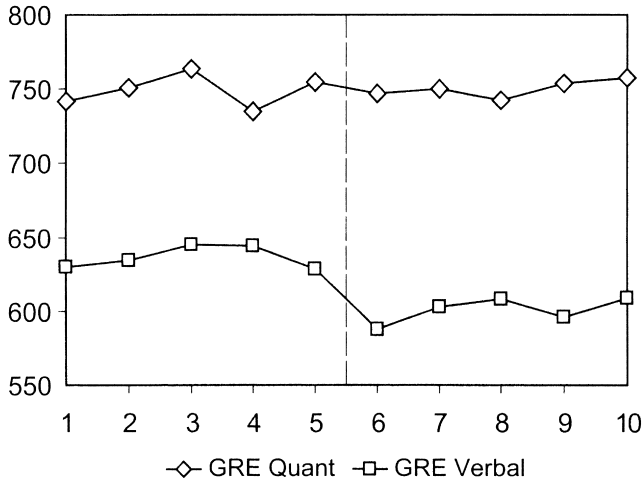
For this scenario, we assume that the goal of fellowship policy is to increase the number of the “best and brightest” pursuing S&E graduate

studies, and that both the numbers and qualifications of GRF awardees contribute to the goal, as described in the model in the appendix. The agency is assumed to face a fixed budget constraint to reflect alternative uses of resources. Our finding that increasing the number of awards reduces average quality while increasing the value of awards raises quality. This yields an optimal solution that depends on the supply response of students to number of awards and to the stipend amount. The critical factor that determines whether it is better to give more awards or raise stipends are the characteristics of applicants who are on the margin of winning awards.

There are two ways to estimate the impact of changes in *awards on the margin* on the measured attributes of awardees. The first is to regress average GREs or other indicators of academic skills on the number of awardees, as we did in table 1.7, where the estimates suggested that average skills were not that sensitive to increased numbers of awards. The second way to assess skills on the margin is to go back to the micro records on individual attributes and rank applicants by the probability that they would get an award and then to examine the characteristics of those who would be given awards on the margin. To do this, we estimated the probability of getting awards in 2004 and then computed the quantitative and verbal GRE scores and GPA for awardees with the lowest probabilities in bins of fifty persons. We also estimated the average GRE and GPA for nonawardees with the highest probabilities of having gained an NSF, again in groups of fifty. The difference in attributes between these nonawardees and the awardees in the lowest group reflects the potential change in quality from increasing the number of awards.

Figure 1.11 displays the results of our calculations in terms of the attributes of awardees and nonawardees. Persons in group 1 are awardees with the highest probability of getting an award. Persons in group 2 are awardees with the next highest probability of getting an award, and so on. Group 5 consists of awardees in the lowest probability group of getting an award. Groups 6 through 10 consist of persons who did not get an award, with those whom we estimated as most likely to get an award placed in group 6, followed by persons with less likelihood of getting an award in groups 7 through 10. The figure shows that there is very little difference in the GRE Quantitative measures around the cutoff point between groups 5 and 6. This indicates that the number of awards could have been increased along a substantial margin without greatly reducing quality, or could have been reduced along a substantial margin without greatly increasing quality. But the figure does show a noticeable drop in GRE Verbal scores between the marginal awardees and the marginal nonawardees. In part, this reflects the fact that while our model does a good job of predicting who gets an award, the errors in the model are for persons along the margin.

Table 1.9 simulates what would happen if NSF has \$10 million addi-



**Fig. 1.11 Quality of GRF applicants on the margin of getting an award**

*Notes:* All persons to the left of the line were given awards. All persons to the right of the line did not get awards. The numbers related to groups of 50 persons, ordered by the estimated probability they would win an NSF award. The 5th group consists of 50 awardees with the lowest probability of getting an award, the 4th group consists of 50 awardees with the next lowest probability, and so on. The 6th group consists of the 50 non-awardees with the highest probability of getting an award. The probabilities are predicted values from an OLS regression of an award receipt dummy variable on panel rating, female dummy, underrepresented minority dummy, and eight field dummies.

tional funding to spend and divided the spending in three different ways: by increasing the number of awards, by increasing the value of awards, and by spending half of the additional money on increasing the number and half on increasing the value. Spending all of the money on increasing the number of awards at current stipend amounts would add 333 awardees and would give awards to persons with modestly lower GRE quantitative and verbal scores than current awardees—4.9 points lower in the quantitative GRE and 29.0 points lower in the verbal GRE. Because NSF gave 1,020 awards, the impact of the lower scores for the 333 awardees on the average for all awardees is considerably smaller. Spending all of the money on increasing stipend values would raise awards by 32.7 percent and would change the overall mean GRE scores of awardees modestly. If such a large increase in stipends were to attract more able persons, this improved quality does not readily show up in the GRE measures. Finally, dividing the 10 million dollars between additional awards and higher valued stipends would increase the number of awards by 146. The increased stipend attracts just enough higher scoring applicants to offset the mean quality decrease associated with more awardees. The result is virtually no change in the average measures of awardee skills.

**Table 1.9** Impact of hypothetical \$10 million increase in annual funding on selected outcomes in 2004, by alternative uses of funds

	Change in annual number of awardees (1)	% Change in relative value of stipend (2)	Difference between mean quality of extra awardees and mean quality of original awardees		Change in overall mean quality of awardees	
			GRE quantitative (3)	GRE verbal (4)	GRE quantitative (5)	GRE verbal (6)
All extra funds to increasing number of awards	333	0.00	-4.9	-29	-1.2	-7.1
All extra funds to increasing value of stipends	0	32.70	n.a.	n.a.	8.2	10
Half to additional awards and half to higher stipend	146	14.30	n.a.	n.a.	3.3	0.7

*Notes:* Nonawardees in 2004 were ranked by their predicted probability of award receipt; see text for details. For a given increase of  $N$  in the number of awards, columns (3) and (4) are calculated by subtracting the mean quality of the actual 2004 awardees from the mean quality of the top  $N$  nonawardees. In column (5)/(6), the change in the overall mean quality induced by more awardees is calculated by multiplying column (3)/(4) by  $N/(1,020 + N)$ , given that 1,020 awards were actually granted in 2004. The change in the overall mean quality induced by higher stipend values is calculated from the coefficients in columns (1) and (3) of table 1.7. The change in overall mean quality induced by both more awards and higher stipend values is the sum of the two preceding changes. n.a. = not applicable.

It is easy to use our analysis to simulate the outcomes of other hypothesized changes in the NSF stipend budget and its allocation between numbers of awards and stipend values.

#### 1.4.2 Using Fixed Rules to Determine the Number and/or Value of Stipends

An alternative way to set the number and/or value of stipends is to use a fixed rule to set them. Monetary economists and officials have often discussed using fixed rules to set the growth of money supply. Some central banks target a range of rate of inflation as their goal while others focus on keeping interest rates within some band. Fixed rules increase social security payments with inflation and have been proposed to raise minimum wages and income tax brackets with inflation rather than through legislative means. If policymakers chose a given adjustment rule, the value or number of stipends would increase according to the rule.

There are diverse rules that policymakers could use to set stipend values and the numbers of stipends. Taking the value of stipends, one possible

rule would be to set stipends as a percentage of alternative earnings in the labor market, such as the earnings of young college graduates. Another would be to make stipends a given percentage of the earnings of doctorates. Yet another would be to adjust stipends to measures of employment opportunities, such as unemployment rates or, given low rates of joblessness among S&E doctorates, the length of time to obtaining a job upon completion of a Ph.D., or to the growth of R&D spending. Since any single indicator would be subject to measurement error, perhaps the most sensible target would be a weighted average of indicators: stipends would rise/fall as the average of the indicators increased or decreased.

To set the number of stipends by a fixed rule, policymakers could simply make them a given proportion of the number of BS graduates, thereby assuring students in different cohorts the same probability of gaining a stipend. When the supply of S&E baccalaureates increased, so too would stipend availability. Alternatively, one could use the same set of weighted indicators for setting the number of stipends as for setting the value of stipends.

To get some sense of how a fixed rule would set the value and number of GRF stipends, we have simulated what would happen to the number of awards and measured scholastic achievement of GRF awardees under two conditions: (a) that NSF fixed the number of awards at 0.41 percent of S&E bachelor's degrees, the level in 1968, before the ratio of awards to bachelor's graduates began to fall; and (b) that NSF set stipends at the 2004 ratio of the value of stipends to the earnings of young college graduates from the current population survey (CPS), 115.8 percent. By setting numbers using a high awards per S&E baccalaureate and setting values at a high value to earnings of young college graduates, we get an upper bound on how a set of fixed rules would affect the supply of applicants to the field. Table 1.10 shows what would have happened from the 1970s to the 2000s if these fixed rules had been employed. The greater number of awards would have lowered measured scholastic skills, but the higher value of awards would have attracted enough applicants with greater skills to allow NSF to raise average quality. Bottom line, these rules would have given awards to a larger number of persons and, given our supply equations, would have raised the average quality of awardees, as well.

Alternative sets of rules would, of course, give different outcomes.

#### 1.4.3 Equilibrating Quality Across Disciplines

Our final policy scenario considers what might happen if NSF changed its charge to panel committees from giving awards in approximately the same proportion to the number of applicants among fields to giving awards in a pure “measured scholastic” achievement, irrespective of the number of applicants from different fields. Thus, if physics received 100 applicants, all of whom had higher GREs, GPAs, and so on than, say social science,



**Table 1.10 Hypothetical GRF program characteristics from maintaining 1968 relative award availability and 2004 relative stipend value**

Period	Actual program characteristics (annual averages)				Hypothetical program characteristics							
	Awards	Awards as % of S&E bachelor's degrees	Average quality of awardees		Change in awardce quality from actual due to more awards		Change in awardce quality from actual due to higher stipend		Predicted quality of awardees			
			Stipend as % of alt. wage	Quantitative	Verbal	Awards	Quantitative	Verbal	Quantitative	Verbal		
1968-1971	921	0.35	55.70	745.2	692.8	1,082	-1.7	-3.5	21.2	25.9	764.7	715.2
1972-1975	511	0.16	66.00	751	693.8	1,285	-9.8	-20.1	16.3	19.9	751	693.8
1976-1979	504	0.17	55.00	753.1	703.6	1,237	-9.5	-19.6	21.6	26.3	765.2	710.4
1980-1983	460	0.15	56.30	746.9	693.1	1,261	-10.7	-22	20.9	25.6	757.1	696.6
1984-1987	520	0.16	82.00	759.6	696.5	1,342	-10.1	-20.7	10	12.2	759.6	688
1988-1991	807	0.25	92.20	766.3	692.9	1,330	-5.3	-10.9	6.6	8.1	767.7	690.1
1992-1995	794	0.22	84.50	762	681.2	1,493	-6.7	-13.8	9.1	11.2	764.4	678.6
1996-1999	820	0.21	72.10	751.1	660.3	1,580	-7	-14.3	13.8	16.8	757.9	662.8
2000-2004	915	0.23	93.40	753.2	633.9	1,640	-6.2	-12.7	6.3	7.6	753.3	628.8

Notes: "Hypothetical" characteristics denote those predicted to prevail if awards as percent of S&E bachelor's degrees had been maintained at 0.41 percent (its 1968 value) and stipend as percent of alternative wage had been maintained at 115.8 percent (its 2004 value). Hypothetical changes in GRE quantitative and GRE verbal are based on coefficients in columns (1) and (3) of Table 1.7. In the calculations, actual numbers of bachelor's degrees in 2002-2004 were imputed from 2000-2001 field-specific growth rates in earned bachelor's degrees, and actual alternative wage in 2004 was imputed from 1983-2003 growth rate in salary of 21-25-year-olds with bachelor's degrees.

which had 500 applicants, all 100 physics applicants would receive awards before any social science applicant would receive an award. To see how awards would be given among disciplines under this hypothetical system, we estimated a linear probability regression of actual award receipt on GRE Quant, GRE Verbal, GPA, reference score (normalized), female indicator, and minority indicator in 1997, the last year for which we had a full set of scholastic qualification measures. In 1997, NSF awarded 850 GRFs, so under the hypothetical system, the applicants with the highest 850 predicted probabilities would be offered an award. Since physical science and mathematics and computer science students have the skills to shift to other fields and do well while students in the social sciences, psychology, and life sciences cannot easily shift into the physical and mathematical sciences, we anticipated that this change in policy would benefit the “harder” sciences.

The calculations summarized in table 1.11 tell a different story. Column (1) gives the number and percentage of applicants in each field in 1997. Column (2) gives the number and percentage of awardees in each field. The percentages of awards by discipline differ moderately from the percentages of applicants by discipline, with engineering gaining more awards and life sciences gaining fewer awards, due in part to the WECS program. Column (3) gives our hypothetical distribution of awards by discipline on the basis of the measured scholastic attributes. The table note shows the coefficients of the equation used to generate the hypothetical distribution. It gives equal weight to the quantitative GRE and verbal GRE and a relatively heavy weight on the measure of references.

The hypothetical distribution shows some differences from the 1997 actual distribution of awardees among natural science fields: mathematics, chemistry, and physics/astronomy gain a few awards while earth and atmospheric sciences and life sciences get fewer awards. The biggest change, however, is between engineering and the social sciences and psychology. Engineering loses a substantial number of awards, where social sciences and psychology gain. Why? The social sciences have the highest mean verbal scores whereas engineering has the lowest mean verbal scores. The social sciences and psychology score best on the reference letters. The GRE quantitative scores favor the physical and mathematical sciences, including engineering, but a larger proportion of students obtain top scores in the GRE quantitative test than in the GRE verbal test, which puts a bound on the extent to which our hypothetical allocation gives awards to those fields. It is possible that other simulations could give a somewhat different story: a high GPA in mathematics or physics may be more difficult to attain than one in the social sciences or biology, and reference letters do not compare students across fields. Still, it would take a very different set of calculations to support our initial expectation: that if NSF gave awards with no reference to fields, the physical sciences and mathematics would get more GRFs and the other sciences would get less awards.

**Table 1.11** Existing and hypothetical allocation of GRF awards across fields in 1997.

Field	Number of candidates (% in parentheses)			Mean applicant quality			
	Applicants	Awardees actual	Awardees hypothetical	Quantitative	Verbal	GPA	Reference
Chemistry	358 (7.0)	57 (6.7)	69 (8.1)	709	586	3.66	-0.05
Computer Science	261 (5.1)	47 (5.5)	46 (5.4)	730	600	3.66	-0.02
Engineering	1,212 (23.6)	252 (29.6)	184 (21.7)	728	573	3.67	0.05
Earth/ Atmospheric	194 (3.8)	32 (3.8)	25 (2.9)	690	603	3.57	-0.05
Life Science	1,502 (29.3)	210 (24.7)	198 (23.3)	686	605	3.54	-0.09
Math	242 (4.7)	41 (4.8)	59 (6.9)	747	614	3.69	-0.12
Physics/ Astronomy	330 (6.4)	58 (6.8)	62 (7.3)	749	632	3.72	-0.26
Psychology	331 (6.4)	48 (5.6)	63 (7.4)	653	615	3.57	0.27
Social Science	698 (13.6)	105 (12.4)	144 (16.9)	667	642	3.60	0.19
Total	5,128 (100.0)	850 (100.0)	850 (100.0)	702	604	3.62	0.00

*Notes:* Awards under the “hypothetical” system are allocated as follows. A linear probability regression of actual award receipt on GRE quantitative, GRE verbal, GPA, reference score (normalized), female indicator, and minority indicator was used to predict the probability of being offered an award in 1997. Under the hypothetical system, applicants with the highest 850 predicted probabilities are designated as being offered an award.

Regression with robust standard errors

Number of obs = 4566

R-squared = 0.1772

gotaward	Coef.	Std. Err.	t
quant/100	.0541323	.0062852	8.61
verbal/100	.0582354	.0056672	10.28
gpa	.1265555	.0153695	8.23
refscore	.0879098	.0050868	17.28
female	.059188	.011009	5.38
minority	.0319515	.0126827	2.52
_cons	-1.04754	.0670683	-15.62

## 1.5 Conclusion

Stipends to U.S. citizens/residents are a natural policy tool for increasing the incentive for Americans to enter S&E fields without directly impacting the flow of talent from overseas. Analysis of NSF Graduate Research Fellowships suggests that raising the value of awards increases the

number of applicants and quality of awardees, while giving more awards increases the number of awardees, by definition, with only a modest reduction in measured academic skills. Because the analysis is limited to a single stipend program, it is uncertain whether the finding that the number of applicants responds to the relative value and relative number of stipends can be extrapolated to the market for graduate students as a whole. To the extent that changes in NSF fellowship policy induce changes in the policies of other stipend-granting groups, we suspect that the qualitative results, at least, can be extrapolated to the broader market.

To see if such an extrapolation is at least consistent with the data, we examined the changing number of first-year first-time graduate students in science and engineering who were U.S. citizens and permanent residents relative to the number of S&E bachelor's graduates in the United States. Appendix tables 1A.1 and 1A.2 give the results of our analysis. Columns (1) and (2) in appendix table 1A.1 show the number and percentage change in all first-year full-time students, including international students. Columns (3) and (4) give the number and percentage change in number of first-year graduate students who were U.S. citizens/permanent residents and thus likely to be affected by higher valued awards for U.S. students. The final columns show an increased proportion of S&E bachelor's graduates enrolling in graduate school over this period. These statistics show greater increase in U.S. citizen/permanent resident first year enrollments during the period when NSF increased the value of awards, and an increase in the number of graduate students relative to S&E bachelor's graduates. The regression analyses in appendix table 1A.2 confirm this reading of the evidence. They relate the logarithm of the number of first-time full-time citizen/residents graduate enrollments to the value of stipends relative to college graduate salaries (column [1]) over the period and the log of the number of enrollments to the number of GRF applicants the previous spring. The estimates in column (1) give an elasticity of enrollments to the value of the stipends of nearly 0.40, which is less than half the elasticity of GRF applicants to stipends found in table 1.6. The estimates in column (2) show that indeed the increased enrollments are associated with increased applicants to the GRF program.<sup>21</sup>

Given the small sample sizes and limited time series variation, we view the appendix tables 1A.1 and 1A.2 analyses as suggestive. The results are consistent with the notion that more and better-paying stipends could raise the number of native-born/residents choosing S&E fields broadly. To go further than that would require a more extensive analysis of the decision of students to pursue S&E graduate studies as opposed to other options (such as simply applying for the GRF) and a careful study of the responses of

21. Analyses that eliminated social science and psychology bachelor's graduates from the calculations gave similar results to those in the appendix tables.

universities, other government agencies, and nonprofit groups to NSF stipend policies.

## Appendix

### *Model of Optimal Stipend Policy*

As a starting point for analyzing stipend policy, we consider the choices and constraints faced by an abstract granting agency. We assume that the granting agency cares about the number of awards ( $N$ ) that it grants each year and the average quality ( $Q$ ) of the students who receive the awards. The real-world analogue is that the NSF would like to support as many students as possible for graduate study in S&E, but at the same time the NSF would like its fellowship recipients to be as high-achieving as possible. The agency thus seeks to maximize

**Table 1A.1** First-time, full-time graduate students in science and engineering

Year	First-time, full-time graduate students in S&E		First-time, full-time graduate students in S&E (U.S. citizens & perm. res.)		U.S. first-time, full-time graduate students as % of U.S. S&E bachelor degrees
	Number of students	Annual growth (%)	Number of students	Annual growth (%)	
1987	71,183	—	48,121	—	15.6
1988	70,894	-0.4	46,478	-3.4	15.2
1989	74,400	4.9	48,309	3.9	15.9
1990	76,597	3.0	49,145	1.7	15.7
1991	81,484	6.4	52,186	6.2	16.1
1992	83,159	2.1	54,603	4.6	15.9
1993	79,366	-4.6	54,027	-1.1	15.3
1994	77,968	-1.8	54,318	0.5	15.0
1995	74,384	-4.6	52,378	-3.6	14.4
1996	73,479	-1.2	51,260	-2.1	13.9
1997	73,544	0.1	50,375	-1.7	13.5
1998	74,431	1.2	49,828	-1.1	13.3
1999	75,447	1.4	48,362	-2.9	12.7
2000	78,332	3.8	46,316	-4.2	12.1
2001	82,411	5.2	48,207	4.1	12.5
2002	86,827	5.4	54,625	13.3	13.7
2003	89,331	2.9	59,649	9.2	14.1
2004	86,565	-3.1	58,853	-1.3	13.5

*Source:* National Science Foundation, Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS), various years; Integrated Postsecondary Education Data System Completions Survey, various years; and authors' tabulations and imputations.

**Table 1A.2** Determinants of first-time, full-time fall graduate enrollment in science and engineering

Independent variables	Dependent variable: Logarithm of first-time, full-time graduate enrollment in S&E by U.S. citizens and permanent residents	
	(1)	(2)
Ln(GRF stipend / outside salary) from previous calendar year	0.397 (0.088)***	—
Ln(GRF applicants) from most recent spring	—	0.284 (0.038)***
Ln(S&E bachelors degrees) from most recent spring	0.462 (0.096)***	0.228 (0.072)***
Observations	18	18
R-squared	0.714	0.854

Notes: Standard errors are in parentheses. Analysis period is 1987 through 2004.

\*\*\*Significant at the 1 percent level.

\*\*Significant at the 5 percent level.

\*Significant at the 10 percent level.

$$(A1) \quad U = aN + (1 - a)Q.$$

a weighted average of  $N$  and  $Q$  where  $a$  is the weight placed on numbers. The agency has two policy levers to do this: the numbers of stipends ( $N$ ) and the annual value of stipends ( $V$ ). The agency faces a budget constraint

$$(A2) \quad NV = K,$$

where  $K$  is the (fixed) total annual budget allocated to stipends.

The agency's choice of  $N$  and  $V$  affects the quality of awardees. Average recipient quality is modeled as

$$(A3) \quad Q = Q(V, N, \text{other factors}),$$

where  $\delta Q / \delta V > 0$  because higher-valued stipends are able to attract higher-achieving candidates who would otherwise have pursued alternative fields and careers or seek other fellowships. We also expect that  $\delta Q / \delta N < 0$ : average quality of the awardees declines with the number of awards because the highest-achieving students are generally the first to obtain awards. The optimal policy in this model, derived from maximizing the objective (A1) subject to the constraints (A2) and (A3), is characterized by the following condition:

$$(A4) \quad a = (1 - a) \left( -\frac{\delta Q}{\delta N} \right) + (1 - a) \frac{\delta Q}{\delta V} \frac{V}{N}.$$

The optimal policy is to equate the marginal benefit and marginal cost of funding another student. The marginal benefit, given on the left-hand side of (A4), is simply the subjective value ( $a$ ) that the agency places on being able to support an additional student. The marginal cost, on the right-hand side, consists of two terms. The first term denotes the marginal decrease in quality resulting directly from giving more awards. The second term reflects the fact that, in the face of a fixed budget, a larger number of awards necessitates a lower annual stipend value, which in turn lowers the average quality of awardees.

This model suggests that the relationship between stipend values and awardee quality is of central interest. How do higher stipend values bring higher-achieving students into the GRFP program? The main mechanism underlying this relationship is that higher stipend values attract a larger applicant pool, and at least some of the additional applicants are highly able candidates who are pulled away from alternative careers (or from immediate entry into the labor market). Essentially, larger applicant pools allow the agency to be more selective in granting awards. Thus, it is also of interest to examine the factors that determine the number of applicants to the GRFP program. We model the number of applicants ( $A$ ) as

$$(A5) \quad A = A(N_{-1}, V, \text{Other factors}),$$

where  $N_{-1}$  is the number of awards granted in the previous year. We predict that  $\delta A / \delta V > 0$  and  $\delta A / \delta N_{-1} > 0$ ; that is, more students apply when stipend values increase and when the perceived probability of receiving an award rises. Other factors predicted to influence the applicant pool include the attractiveness of alternative careers.

Thus, equations (A1) through (A5) capture the quality-quantity choice facing a stipend-granting agency: to spend its budget on funding many students versus funding a few “superstars.” Since identification of stars is difficult in most disciplines, the choice is in fact more complex. The key behavioral relations to be empirically estimated are: the change in the quality of students as the number of awardees is increased; the change in quality in response to the value of stipends; and the change in the number of applicants as stipend values rise. If the agencies responsible for stipend policy optimizes as in this model, then the preceding model reflects the basic tradeoffs factoring into the agencies’ decisions. In this case the number and monetary value of GRFP fellowships cannot be assumed to be exogenous for the purposes of estimating behavioral relationships. On the other hand, if the number and value of GRFP fellowships depend on vagaries of budgeting, then adjustments to the number and value of stipends constitute “natural experiments” for assessing the effect of stipend policy.

## References

- Committee of Visitors (COV). 1999. Report of the Committee of Visitors, National Science Foundation, The Graduate Fellowship Program. June 17–18.
- . 2003. Committee of Visitors (COV) report for the National Science Foundation (NSF) Graduate Research Fellowship Program (GRFP). June 17–18.
- Freeman, R. B. 2005. Fellowship stipend support and the supply of S&E students: NSF graduate research fellowships. Paper presented at Annual Meeting of the American Economic Association. 7–9 January, Philadelphia, PA.
- Goldsmith, S., J. Presley, and E. Cooley. 2002. *National Science Foundation graduate research fellowship program: Final evaluation report*. Arlington, VA: National Science Foundation.
- National Science Board. 2004. *Science and engineering indicators*, vol. 2., appendix table 2-15. VA: National Science Foundation.
- National Science Foundation, Division of graduate education. 2002. Guideline for panelists in the review and rating of National Science Foundation Graduate Research Fellowship applications. Arlington, VA: National Science Foundation.
- National Science Foundation, Graduate Research Fellowship Program (GRFP). 2007. Program solicitation, NSF 07-576. Arlington, VA: National Science Foundation.
- National Science Foundation, Division of Science Resource Statistics. 2001. Graduate students and postdoctorates in science and engineering: Fall 2001 tables 39, 40. Arlington, VA: National Science Foundation, Division of Science Resource Statistics.
- Roy, A. D. 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers* 3 (2): 135–46.