Science and Public Policy, 2017, 1–21
doi: 10.1093/scipol/scx017
Article

Influences of academic institutional factors on R&D funding for graduate students

Alexandra Graddy-Reed[1,*,†], Lauren Lanahan[2,†] and Nicole M. V. Ross[3]

[1]Price School of Public Policy, University of Southern California, 650 Childs Way, Los Angeles, CA 90089, USA, [2]Lundquist College of Business, University of Oregon, 1208 University St, Eugene, OR 97403, USA and [3]Department of Public Policy, University of North Carolina at Chapel Hill, Abernethy Hall, C.B. 3435, Chapel Hill, NC 27599, USA

*Corresponding author. Email: graddyre@price.usc.edu
†The first two authors contributed equally to this article and are considered joint first authors.

Abstract

We examine the effect of academic institutional characteristics on research funding grant success for graduate students. This article draws upon the US National Science Foundation's Graduate Research Fellowship Program (GRFP). We match a set of graduate students to their graduate programs to examine whether higher education institutional factors mediate funding assignment to award or honorable mention. We find evidence that a series of leadership, peer, programmatic, and university characteristics are associated with grant funding outcomes. Notably, faculty research and peer quality are associated with award success, while the signal of being at a public institution decreases the likelihood of award receipt. Moreover, while we find that larger programs are more likely to have graduate students that receive awards, the larger, lower-ranked programs exhibit inefficiencies in scaling the activity. This implies that these programs may face coordination costs that are detrimental to the rate of graduate student success.

Key words: R&D funding; higher education; graduate students; science and engineering; federal funding.

1. Introduction

US universities compete annually for over $63 billion in science and engineering (S&E) research and development (R&D).[1] As a result, considerable research has studied the broad implications of these investments on the research enterprise. The literature to date has focused primarily on research production among senior scholars, teams, and labs (e.g. Zucker and Darby (1996); Bozeman and Corley (2004); Azoulay et al. (2007); Bercovitz and Feldman (2008); Jones et al. (2008); Conti and Liu (2015); Lane et al. (2015)). Recent efforts have expanded the focus to examine how funding schemes specifically affect high-risk/high-reward research, with attention to junior faculty (Heinze 2008). However, less attention has focused on the research environment and funding schemes for graduate students.

In this article, we focus on the large, yet understudied group of S&E graduate students—referred to henceforth as emerging researchers. These individuals are at the beginning of their careers and receive formal and informal training from their graduate programs as they establish and build their research acumen. While emerging researchers often contribute to their graduate programs as research assistants, they are also expected to develop and pursue their own research agenda.

We examine the role of higher education institutions on emerging researcher success given the importance of the academic environment at this early stage. Initial institutional conditions are likely to have persistent 'imprinting' effects on future productivity (Stinchcombe 1965). Not only does the academic setting provide the physical infrastructure for research production, graduate programs also serve as the foundational base for knowledge transfer through course offerings, apprenticeships with faculty, research seminars, and more informal networking opportunities (Bandura 1986; Bercovitz and Feldman 2008; Cetina 2009; Azoulay et al. 2011; Stephan 2012; Agarwal and Ohyama 2013). Nonetheless, as emphasized by Lane and Bertuzzi (2011), empirical studies within this academic context are limited (Lane and Bertuzzi 2011). In response to a series of recent discussions in *Nature* lamenting our nascent understanding of S&E graduate programs (Callier and Polka 2015; Gould 2015; Woolston 2015; Nature 2015), we direct attention to the academic setting centered at this early career stage.

We focus on external grant funding receipt as an early stage milestone. In developing a research pipeline, funding attainment demarcates not only the initiation of a research line, but also signals professional promise for the principal investigator (PI) (Stinchcombe 1965). Within the S&E fields, resource acquisition is a necessary precursor for research production that includes outputs like academic publications and patents. Moreover, the constant threat of the ever-diminishing public research funding[2] compounds this issue.

© The Author 2017. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

The US National Science Foundation's (NSF) Graduate Research Fellowship Program (GRFP) provides a rare opportunity for emerging researchers to obtain a substantial and prestigious research award. The grant is unique and relatively sizeable in value, providing the most meritorious students with three years of guaranteed funding to pursue their own research agenda with no service obligation to their graduate program.

For this study, we draw upon a large sample of GRFP awardees and honorable mentions to identify a comparable set of S&E graduate programs. From this set, we assess what institutional characteristics (program- and university-level)[3] are associated with students obtaining external R&D funding. We focus on the institution at this early career stage as it offers legitimacy to emerging researchers who aim to alleviate uncertainty given the novelty of their research pipeline and their nascent research record (Stinchcombe 1965; Aldrich and Fiol 1994).

The economic and educational implications of this study are notable. The US National Science Board estimates that in 2013 approximately 615,000 individuals enrolled in US-based S&E graduate programs.[4] The economic impact of this population is not only defined by the sheer supply of graduate students, but also by the productive research endeavors they engage in (Conti and Liu 2015). From an educational standpoint, graduate programs not only house a substantial portion of the research enterprise, they train the next generation of researchers. This study offers a large-scale empirical assessment of the relationship between higher education institutional factors and emerging researcher outcomes.

The article proceeds as follows: section two presents the institutional framework, highlighting the implications of higher education institutional features as mediators for emerging researcher R&D grant receipt. Section three presents the research design detailing the data, sample, and methods. Section four presents the results while section five provides a discussion and concluding remarks.

2. Mediators of R&D: institutional factors

As 'natural incubators', academic programs and their respective university enhance a relatively steady stream of researchers' activity with human and physical capital (Etzkowitz 2003). Academic programs encourage research in part by giving autonomy to its personnel, allowing researchers to experiment and push research boundaries forward (Amar et al. 2009). Given the growing financial costs to conduct research, R&D resources play a central role in facilitating the research enterprise (Stephan 2012). Yet, little attention has been directed toward understanding the role of graduate programs on emerging researcher funding attainment.

Given the early-stage of an emerging researcher's career, the role of institutional factors may be more significant in the acquisition of research funding. For one, institutional characteristics may impact receipt of funding by offering a more attractive research environment to prospective high-quality students. Second, graduate programs may also provide students with a higher level of training or improved access to research support. Third, given the strength of professional networks across academic disciplines (Friedman and Friedman 1982; DiMaggio and Powell 1983), the reputation of the graduate program can offer positive signals to the broader academic community. Moreover, the reputation of the institution likely alleviates uncertainty around the emerging researcher's abilities given both the student's early career stage and abbreviated nature of the basic science research proposal. Of note, this influence may stem from multiple levels within the higher education institution, spanning the academic program within which the emerging researcher is enrolled to the larger university.[5] Altogether, the academic institution can have direct impact not only on external resource acquisition, but also for later stage milestones as measured by academic publishing and professional placement. Below we review a series of factors across multiple levels of the higher education institution that may mediate funding acquisition through any of these mechanisms. These include leadership, peer, programmatic, and university characteristics.

Faculty leadership is arguably most likely to impact the success of graduate students given the emphasis on apprenticeships within academia. Students observe the actions of their leaders and learn what activities are deemed legitimate (Bercovitz and Feldman 2008). Through a process of social learning, students then emulate the behaviors considered most appropriate and even adopt value systems that resemble those of their mentors (Shamir et al. 1993). Recent research has found evidence of this through co-authored publication trends between graduate students and their academic advisors (Pinheiro et al. 2014). In addition, faculty with high research and publication activity from the applicant's graduate program can create the potential for knowledge spillovers between faculty and students. Moreover, the reputation of the faculty may be prominent across the discipline and thus familiar to a reviewing panelist. The human capital gained from a productive faculty could impact the probability of award obtainment through any of these mechanisms.

Peer productivity can have similar effects. Bercovitz and Feldman (2008) find that peer groups act in tandem with leaders and serve as an important reference (Bercovitz and Feldman 2008). In graduate training, students often look to the behaviors of other graduate students who share similar research interests or experiences as inspiration for how to approach problems and make decisions (Bandura 1986; Ellison and Fudenberg 1993; Duflo and Saez 2000; Sorensen 2002). Within an academic program, this means that actions taken by one student hold important spillover effects for their peers. For example, programs with a high volume of graduate research assistants may spur more productive research output than programs with a higher volume of graduate teaching assistants. A student's research experience may inspire other students to seek out similarly prestigious and beneficial research opportunities. This is important because, as recent empirical work suggests, collaboration—as opposed to isolated production—has become an increasingly important element in research output (e.g. Pinheiro et al. (2014); Aggarwal et al. (2015); Wuchty et al. (2007); Singh and Fleming (2010)).

From another angle, diversity may spur knowledge spillovers; this has the ability to foster productivity, research agendas, and perspectives (Amin and Cohendet 2000; Autio et al. 2014). Organizations can be diverse in a multitude of ways based on the composition of their personnel by spanning a multitude of technical skill sets, experience levels, normative ideologies, and general demographic characteristics. In the case of technical skill sets, if the program fosters interdisciplinary interaction (as opposed to only within-discipline interactions), students are likely to draw upon a broader set of skills and analytical perspectives for approaching a research problem (Lattuca 2001). Critics to this argument, however, highlight that team diversity can also lead to higher coordination costs (Williams and O'Reilly 1998; Aggarwal et al. 2015). Within an academic program, these costs could arise with interdisciplinary research agendas and even in instances when researchers from different backgrounds interact (Jacobs and Frickel 2009).

Programmatic resources can also serve a critical role in connecting human and physical capital with one another. Recent research has noted the importance of comprehensive support mechanisms for graduate students as they pursue a variety of diverse career paths (Agarwal and Ohyama 2013). Within the higher education literature, others have noted the significance of intradepartmental policies and practices in encouraging graduate student persistence and performance (Nerad and Cerny 1993; de Valero 2001). The evidence suggests that regular evaluations of student progress and an emphasis on apprenticeship and collaboration leads to improved student outcomes.

Finally, we argue that the university may mediate the grant application outcome. While at a higher organizational level and more distant to the student than the graduate program, we argue that university ranking and the type of institutional control not only signal research quality, but also the relative capacity to provide resource support. In particular, a relatively nascent stream of literature has focused on the role of institutional control on research output finding private institutions to be more equipped to secure and support research than public universities (Aghion et al. 2010; Whalley and Hicks 2014).

3. Research design

The premise of this analysis is to assess what programmatic and university characteristics are associated with receipt of a prestigious research award that is designated for emerging researchers. These organizational factors may impact award receipt through human and physical capital advantages by improving the training, offering research support, or providing a positive quality signal to proposal reviewers.

High-quality students are likely attracted to programs and universities that provide competitive research support. Thus, while there is an endogeneity concern that high-quality students select into high-quality programs and universities, we argue that institutional characteristics also mediate the funding process at this early career stage. To address the concern of endogeneity, we follow a line of literature that focuses around the funding cutoff (Arora and Gambardella 2005; Goldfarb 2008; Azoulay et al. 2011; Jacob and Lefgren 2011a,b). Rather than sampling from the full population of graduate students programs, we alleviate baseline concerns of endogeneity by drawing upon emerging researcher grant proposal data that reports both awardees and honorable mentions. The latter are competitive applicants who are recommended by proposal reviewers, but who are just shy of receiving the funding. Without granular data on research quality metrics for the student-PIs, we recognize that student quality remains an empirical concern. As such, we do not purport to definitively identify a counterfactual; the interpretations of the results are thus associative.

Critical to this design is the fact that reviewing panelists distinguished both groups of applicants—awardees and honorable mentions—as having high merit and research potential. This designation defines the set of graduate programs and provides a baseline of student quality. Importantly, these programs not only contain emerging researchers who actively seek graduate research funding, but also received some form of formal recognition.

To strengthen the design, we use doctoral program rankings to further proxy for the innate student quality of the application. Specifically, we match comparably ranked doctoral programs that only have graduate student honorable mentions to those with any award activity. Effectively, we aim to define similar subsamples of graduate programs with comparable prestige and quality to assess moderated effects of program rank.

This section proceeds as follows. First we present an overview of the data, sample, and empirical methods. Then in building upon the discussion directly above, we present a series of stratification techniques to strengthen the primary analysis. Finally, we discuss the functional forms of the variables and present the descriptive statistics.

3.1 Data

We use data from two separate databases: the NSF's GRFP grant database and the National Research Council's (NRC) data on US research doctorate programs.

3.1.1 GRFP proposal data

The GRFP has a demonstrated history of supporting promising S&E graduate students. Award recipients receive a generous three-year fellowship to conduct their own research. In 2015, NSF offered $138,000 for the full award—$34,000 as an annual student stipend and $12,000 as an annual educational allowance to the institution.[6] The following features make this funding mechanism particularly salient to this analysis. First, the program is designed specifically to support emerging researchers; applicants are only eligible prior to starting a graduate program up through the fall of their second year of graduate school. Given this study's focus on graduate programs as an antecedent on funding receipt, we sample graduate students from the GRFP database who were currently enrolled in a graduate program at the time of GRFP recognition.[7] Second, as a signal of quality, both award recipients *and* honorable mentions are publicized, collectively representing the top 20 percent of all student applicants. Panels—composed of senior disciplinary and interdisciplinary scholars—review applications and make recommendations to NSF regarding who should receive the award. NSF then 'determines the successful applicants from these recommendations, with Fellowships and Honorable Mentions offered based on the GRFP portfolio within the context of NSF's mission'.[8] While awardees receive full funding, the acknowledgement of the honorable mention is revered as a signal of intellectual merit and research promise.

While the single-blinded merit review component for this program mirrors other NSF proposal reviews, the GRFP application notably differs. Standard, collaborative, career, dissertation, workshop, and larger center NSF proposals are all approximately three times longer in length allowing for greater explication of research proposal. NSF has tailored the GRFP application requirements given the nature of the funding opportunity for emerging researchers.

To our knowledge, only a handful of studies have used the GRFP database (Chapman and McCauley 1993; Bartolone et al. 2014; Le and Bartolone 2015). While this set of studies focuses primarily on programmatic outcomes of the award, we offer a more foundational analysis by focusing on how institutional factors mediate these outcomes.

The GRFP data is a time series dataset structured at the student-proposal level and includes institutional affiliation and field of study for the population of awardees and honorable mentions. For the analysis we compute aggregated annual, program-level counts, indexed by academic field, i, and university, n, of student awardees and honorable mentions, respectively.

3.1.2 NRC programmatic data

The NRC is a decennial survey assessment of the quality of US research doctoral programs (Ostriker et al. 2011). The most recent survey (2005–6), published in 2010, offers the most granular,[9] rigorous, and extensive assessment of graduate program quality containing detailed information on program characteristics for a representative sample of doctoral programs (Hicks 2009; Schmitt 2013). This survey includes data on a series of program-level measures ranging from faculty publications, citations, grants, and diversity to characteristics of the graduate student population including average Graduate Record Examination (GRE) scores, type of financial support, and sociodemographic characteristics. The data also contain information on characteristics of the graduate program such as the number of PhD's granted over five years, median time to degree, student completion rate, and various types of graduate student support activities. Over 5,000 doctoral programs that span 62 academic fields from 212 universities were surveyed (see Chapter 3 and Appendix D of Ostriker et al. (2011)).

Other studies that draw upon this data have placed precedence on program review to promote academic standards. Moreover, much of the scholastic attention has focused on discussions of higher education ranking methodology (e.g. Schmitt 2013; Brooks 2005; Dill 2006; Dill and Beerkens 2013). Less attention, however, has been directed toward utilizing the data for higher education empirical analyses despite the richness of the data.

Taken together, we merge the GRFP and NRC databases based on a numeric university-field crosswalk. We refer to publicly available National Center for Education Statistics (NCES) Integrated Postsecondary Education Data System (IPEDS) data and the taxonomy of NRC's S&E fields to assign a unique university and field identification, respectively. We draw upon the individual's listed field and university from the GRFP database to identify the program, and then match to the NRC data for detailed graduate program characteristics. The level of analysis for this study is at the graduate program, indexed by the academic field, i, and university, n. The Appendix details the data building process.

3.2 Sample

The NRC programs with *any* GRFP award and/or honorable mention activity by students from 2005 to 2008 define the sample. We delimit the sample by graduate programs with at least one GRFP award or honorable mention to maintain a more representative set of observations. Given the S&E scope of the GRFP program, this includes NRC programs from 41 academic fields from four broad divisions of Engineering, Life Sciences, Math and Physical Sciences, and Social and Behavioral Sciences, as defined by the US National Academy's Board on Higher Education and Workforce.

The GRFP reports annual data on proposal activity while the NRC program-level data is a cross section. Although administered as a decennial survey, the NRC data reflect programmatic trends over several years. The most recent 2010 survey contains data from years 2000 to 2006. Thus, we draw upon GRFP data over a four-year period, 2005 to 2008, notably a timeframe that overlaps the tail end of NRC data collection efforts. We exclude the years after 2008 where the Great Recession had significant impact on the economy, affecting both NSF funding and GRFP activity. Of note, the number of GRFP awards roughly doubled from 2009 to 2010;[10] moreover, the ratio of honorable mentions to awards significantly decreased.[11]

Table 1. Distribution of GRFP activity for full sample

	Full sample
Graduate programs with GRFP activity	1,033
Number of academic fields	41
Number of universities	142
Total number of awards and honorable mentions	5,626
Program mean count	5.45
GRFP awards	
Number of programs with awards	643
Share of programs with awards	62%
Total number of awards	1,984
Program mean count	3.09
GRFP honorable mentions	
Number of programs with honorable mentions	906
Share of programs with honorable mentions	88%
Total number of honorable mentions	3,642
Program mean count	4.02

Note: Sample includes only graduate programs with GRFP activity (award and/or honorable mention) between 2005 and 2008.

Our full sample includes 1,033 unique S&E graduate programs (20 percent of the NRC full sample) from 142 US universities. This sample of programs secured 1,984 awards and 3,642 honorable mentions, for a total of 5,626 GRFP accolades from 2005 to 2008. The average number of GRFP acknowledgements (awardees and/or honorable mentions) in a program over this time period is 5.45. Table 1 provides descriptive statistics of GRFP activity for the sample of graduate programs in the analysis.

3.3 Methods

To examine the effect of institutional factors on funding receipt for emerging researchers, we estimate two empirical models at the program level of analysis.

Step 1: Any Award Success

First, we examine the effect of a set of institutional factors on the difference between a program containing students that win honorable mentions *exclusively* and those that contain students who win at least one formal award (Equation (1)).[12] From the sample of programs with any GRFP activity, 38 percent (390 programs) have students who only obtain honorable mentions, with the remaining 62 percent (643) containing at least one student in receipt of a GRFP award.[13]

$$\begin{aligned}&\Pr(\text{Any GRFP Award}_{in})\\&= f(\beta_0 + \beta_1\text{Leadership Quality \& Composition}_{in}\\&\quad +\beta_2\text{Peer Quality \& Composition}_{in}\\&\quad +\beta_3\text{Program Support \& Traits}_{in}\\&\quad +\beta_4\text{University Traits}_n + \lambda_i + \varepsilon_{in})\end{aligned} \tag{1}$$

The binary outcome of interest is any GRFP award for students in the graduate program from 2005 to 2008, where one indicates that at least one award was received, and a zero indicates that only honorable mentions were received during this time period.

The set of higher education institutional factors are divided into four categories. The first, *Leadership Quality and Composition*, is measured with a vector of variables, β_1, to capture the graduate program's faculty-based characteristics. These include faculty publications and external grant activity to proxy for faculty quality, as well as measures of the share of female, interdisciplinary, and non-Asian minority faculty. We include a similar set of graduate student measures for *Peer Quality and Composition*, β_2. These include the average GRE quantitative scores to proxy for student quality, and measures of the share of female, non-Asian minority students, and the percent of students with academic plans. *Program Support and Traits* are estimated with a vector, β_3, which includes whether there is student workspace, proposal support, and an above average number of student support programs;[14] it also includes median time to degree and the program size by quartile ranking (with the smallest bin as the referent). Program size quartile groupings are based on the entire sample of NRC programs from the original survey. *University Traits* are included with β_4; this includes a vector containing the region, whether it is a public institution, and university rank by tercile. The rank tercile groupings are based on Barron's university rankings with the low bin—formally representing Competitive and Less Competitive institutions—as the referent.[15]

Finally, we include a field fixed effect (λ_i). This refers to the 41 S&E fields from the NRC study. Inclusion of academic field fixed effects controls for unobserved time-invariant variation across academic disciplines and therefore offers more conservative estimates of the models. This allows us to control for differences in academic culture, environment, and structure between fields, even within the same broad divisions (Gardner 2009). The complete list and frequency of fields are provided in Appendix Table A.1.

Step 2: Concentration of Awards

We next use the same sample to assess how these characteristics impact the *share* of students in a graduate program that receive an award (Equation (2)). This continuous variable stands in contrast to the binary outcome presented in Step 1.

$$\begin{aligned}(\text{Awards}_{in}/\text{Eligible Cohort}_{in}) \\ = f(\beta_0 + \beta_1\text{Leadership Quality \& Composition}_{in} \\ +\beta_2\text{Peer Quality \& Composition}_{in} \\ +\beta_3\text{Program Support \& Traits}_{in} \\ +\beta_4\text{University Traits}_n + \lambda_i + \varepsilon_{in})\end{aligned} \tag{2}$$

The continuous outcome variable accounts for the relative concentration of award activity. The variable represents the share of the average eligible cohort that receives the GRFP award contingent on the program having any GRFP activity. The numerator is the average annual award count and the denominator is the average first-year cohort size reported in the NRC multiplied by two. The cohort size is doubled due to the fact that students are eligible in their first and second year of graduate school. Because receipt of a GRFP is a relatively rare event, the average share of award activity is quite small at 2.1 percent with a range of 0 to 42 percent. Standing as an outlier, Oregon State University's Animal Science Program leads the funding activity with 42 percent of eligible graduate students receiving the GRFP award between 2005 and 2008.

We maintain the relative activity of this outcome measure, but normalize the distribution by standardizing the variable so that it has a mean of 0 and a standard deviation of 1, $[(y_{in} - \bar{y})/\widehat{\sigma}]$. Moreover, the variable is standardized *within* the corresponding academic division the program belongs too (refer to Appendix Table A.2) to account for variation across broad academic divisions. Thus, we are estimating how these factors impact the relative concentration of awards. The same set of vectors used in Step 1 are used in Step 2.

3.4 Stratification by program quality

In addition to estimating the primary models with the full sample, we estimate Equations (1) and (2) with stratified samples by program rank. We use the R ranking produced by the NRC—a regression-based ranking derived from a faculty survey of peer programs (Ostriker et al. 2011). In doing so, we match programs of comparable rank to address the potential for confounding factors related to program quality that might attract higher quality students (Hegde 2005). Moreover, what matters for top rank programs may be different from what matters for lower-ranked programs that will not be competing for the same group of students. This approach measures the moderated effect of program quality. We estimate two bins from the tercile rankings—rank 1 (high) and rank 2 and 3 (mid and low). We combined the lower two terciles given the small sample sizes.

3.5 Variable functional form

In each equation, four vectors of variables are included to estimate the role of institutional spillover factors on a program having students to obtain a GRFP award. Appendix Table A.3 details the variables in each vector. To ease interpretation and compare the relative influence of each variable, all continuous variables are standardized so they have a normal distribution with a mean of 0 and a standard deviation of 1, $[(x_{in1} - \bar{x}_1)/\widehat{\sigma}_1]$, where the slope coefficient is $\left(\frac{\widehat{\sigma}_1}{\sigma_y}\right)\widehat{\beta}_1$. This allows us to compare effects across the regressors to assess the probability of having an award (Step 1) or the relative concentration of awards (Step 2) from a one standard deviation increase of each continuous variable.

As with the award concentration outcome variable (Step 2), each variable is standardized with respect to the program's corresponding academic division to account for academic divisional differences.[16] While programs within a university often adopt similar policies and norms, there is a stronger convergence among programs that are across universities within the same academic field (Friedman and Friedman 1982). Programs in the same field residing at different higher education institutions compete across the discipline over students, faculty, funding, and publications to gain legitimacy and prestige.

3.6 Descriptive statistics

Table 2 presents the descriptive statistics of the GRFP performance outcome variables and covariates for both the full set of graduate programs and then subsequent stratified subsamples by program rank with notation on statistically significant differences. Of note, Table 2 reports the baseline statistics; however, we include the standardized measures in the set of regressions for comparative purposes when interpreting marginal effects.

On average 62 percent of programs in the sample have any GRFP award activity with higher levels in rank 1 (71 percent) versus ranks 2 and 3 programs (50 percent), significant at the 1 percent level. Moreover, approximately 2 percent of eligible graduate students receive the prestigious GRFP award. Turning to the set of institutional factors, the average annual number of publications per faculty is 1.78 and 78 percent have research grants. We find higher levels of research output in the rank 1 graduate programs in contrast to ranks 2 and 3 programs (both statistically significant at the 1 percent level). As for faculty composition, roughly 20, 26, and 4 percent comprise female, interdisciplinary, and non-Asian minority faculty, respectively. For peer-related factors, the average GRE quantitative score is 727 (out of a possible 800) with the higher

Table 2. Descriptive statistics of covariates

	Full sample	Rank 1	Ranks 2 and 3	
GRFP performance				
Departments receiving any awards	0.62	0.71	0.50	***
Award concentration	0.02	0.03	0.02	***
Range: 0–0.42	(0.04)	(0.04)	(0.03)	
Award concentration given award	0.04	0.04	0.03	
Range: 0.001–0.42	(0.04)	(0.04)	(0.04)	
Leadership quality and composition				
Average publications per faculty	1.78	2.09	1.34	***
Range: 0.01–10.16	(1.28)	(1.42)	(0.88)	
Faculty with grants (0–100) %	78.13	80.30	74.99	***
Female faculty (0–100) %	20.27	20.40	20.09	
Interdisciplinary faculty (0–95) %	26.16	27.43	24.32	**
Non-Asian minority faculty (0–55) %	4.38	4.21	4.62	*
Peer quality and composition				
Average GRE quantitative score	726.68	740.10	707.24	***
Range: 485.83–800	(51.93)	(43.50)	(56.82)	
Female students (0–91) %	41.31	42.01	40.28	*
Non-Asian minority students (0–100) %	10.79	10.18	11.67	***
Students with academic plans (0–100) %	55.32	57.67	51.91	***
Department support and traits				
Student workspace provided	0.80	0.80	0.81	
Student proposal support provided	0.69	0.70	0.68	
Number of student programs	16.50	16.74	16.14	***
Range: 0–18	(1.79)	(1.51)	(2.08)	
Median time to degree, years	5.59	5.73	5.38	***
Range: 2–12	(1.05)	(0.99)	(1.10)	
Program size quartiles				
Q1, smallest	0.09	0.03	0.19	***
Q2	0.21	0.13	0.33	***
Q3	0.29	0.27	0.33	**
Q4, largest	0.40	0.57	0.16	***
University traits				
Region				
Northeast	0.25	0.28	0.20	***
Midwest	0.22	0.22	0.22	
South Atlantic	0.17	0.15	0.20	**
South Central	0.10	0.06	0.14	***
West	0.27	0.29	0.23	**
Public university	0.65	0.56	0.79	***
University rank				
High	0.66	0.79	0.48	***
Mid	0.26	0.19	0.36	***
Low	0.08	0.03	0.16	***
Observations	1033	611	422	

Notes: Means or proportions for baseline measures are presented; standard deviations are in parentheses; stratification by program rank; Table A.3 details variable composition. Statistical significance based on *t*-tests by program rank. $^{***}P < 0.01$, $^{**}P < 0.05$, $^{*}P < 0.1$.

ranked programs exhibiting higher scores (significant at the 1 percent level). The distributions of minority graduate students are higher than the distributions of faculty; this follows in line with recent statistics that report greater diversity among earlier stage cohorts in higher education (Organization for Economic Cooperation and Development 2012).

For program features, the length of time to complete the degree for rank 1 programs is 5.73 years, which exceeds the ranks 2 and 3 programs by roughly four months (significant at the 1 percent level). As for program size, program rank is positively correlated with program size. Ranks 2 and 3 programs, on the other hand, exhibit a more standard distribution with a larger share of programs with middle-sized programs. As for region, over a quarter of the programs are located each in the Northeast and West. Finally, roughly two-thirds of the programs are within public universities, and the majority of programs are housed within high-ranked institutions according to the Barron's ranking. This distribution mirrors program rank as well with higher ranked programs located at higher ranked universities (significant at the 1 percent level).

To assess the comparability between the group of programs with only honorable mentions to those with at least one GRFP award, we also estimated the comparison of means between the two samples. We find that programs with at least one award have a higher level of average faculty publications and grants but a lower share of female faculty. Programs with at least one awardee have a higher average quantitative GRE score, are likely to be larger in size, and nested within highly ranked, private universities. It is worth noting that the differences for most variables are economically insignificant. This is reported in the Appendix, Table 4.

4. Results

Results are presented below for each step—the logistic estimation of Equation (1) in Section 4.1 and the OLS estimation of Equation (2) in Section 4.2.

4.1 Step 1: Any award activity

We estimate Equation (1), a binary model of any GRFP award success on academic institutional characteristics. Table 3 presents the average marginal effects from the logistic regressions for the full sample in column 1 and the stratified samples by program rank (rank 1 in column 2 and ranks 2 and 3 in column 3). The results for the standardized continuous regressors are interpreted such that a one standard deviation increase is associated with a change in the probability of having at least one student in the program secure GRFP award recognition from 2005 to 2008. The binary regressors are interpreted as the differential effect in the probability of having any award success.

In the results for the full sample (column 1), the following regressors are statistically significant: standardized average number of publications per faculty; standardized percent of interdisciplinary faculty; standardized average GRE quantitative score; binary indicators for the number of student support programs, program size (Q3 and Q4, both with reference to Q1), and the binary indicator for public university. Normalized continuous regressors offer easier comparison of the marginal effects within each model.

Notably, increasing the average number of publications per faculty by one standard deviation is associated with a 7.2 percentage point increase in the probability of the program having any GRFP award activity. The marginal effects for a standard deviation increase for the share of interdisciplinary faculty and the average GRE score are associated with a 3.0 and 6.7 percentage point increase, respectively, in the probability of the program having any GRFP award activity. In other words, programs with greater levels of faculty publication activity, interdisciplinary faculty, and students with higher GRE scores are positively associated with GRFP award activity, holding constant the academic field of the program.

Turning to the set of binary regressors, the differential effect of having an above-field-average number of student support activities is associated with an 8.3 percentage point increase in having at least one student successfully obtain a GRFP award. In other words, programs with a higher proportion of student support activities such as travel funding, teacher training, or proposal support than the average program in their narrow academic field (e.g. economics) are more likely to have student applicants that achieve the GRFP award. Moreover, larger programs have a positive association compared with the smallest programs (Q1). Relative to Q1 programs (the reference group) Q4 programs have a 22.0 percentage point advantage in likelihood of GRFP award activity; additionally, Q3 programs also have a higher probability of award receipt, though the marginal effect is smaller at 9.6 percentage points. Notably, when it comes to university traits, we estimate a large negative effect for public universities. The differential effect of the doctoral program belonging to a public—in contrast to a private—university, is associated with a 14.8 percentage point decrease in the probability of having any GRFP award activity.

When stratified by program ranking, results vary across rank 1 (column 2) and ranks 2 and 3 (column 3). The same set of significant regressors are robust and larger (in absolute value) for the rank 1 sample compared with the full sample, indicating that highly ranked doctoral programs moderate the average effect. Additionally, all three size-related program dummies are positive and significant (again, results are in reference to Q1, the smallest program size). Standardized measures of average faculty publications and percent interdisciplinary faculty, along with binary measures of above-field-average student support activities and public university, are only robust for top-ranked programs. These regressors are not significant for lower-ranked programs (ranks 2 and 3). For the sample of programs with lower ranks, the measure for students with academic plans is significant, where a one standard deviation increase is associated with a 4.6 percentage point increase in the probability of having at least one applicant receive a GRFP award.

4.2 Step 2 award concentration

Step 2 includes a standardized continuous outcome measure to account for the relative share of GRFP awards to eligible cohort size within a program. We interpret the coefficients for the set of standardized regressors as follows: if x_1 increases by one standard deviation, then the share of GRFP awards to eligible cohort changes by β standard deviations. The set of binary regressors are interpreted as follows: the differential effect of x_2 is associated with a β standard deviation change in the share of GRFP awards to eligible cohort. Table 4 presents the coefficients from the OLS regressions for the full sample in column 1 and the stratified sample by program rank (rank 1 in column 2 and ranks 2 and 3 in column 3).

For the full sample, the following regressors are statistically significant: standardized average number of publications per faculty, standardized average GRE quantitative score, program size (Q3 and Q4, both in reference to Q1), the binary indicator for public university, and dummy for high-ranked university (in reference to low). Increasing the average number of publications per faculty (average GRE quantitative score) by one standard deviation is associated with a 0.151 (0.159) standard deviation increase in GRFP award concentration. Moreover, the differential effect of high university rank in reference to low is associated with a 0.203 standard deviation increase in GRFP award concentration. Being a public university, meanwhile, as opposed to a private university is associated with a 0.186 standard deviation decrease in GRFP award concentration.

Finally, the results for program-size are negative, in contrast to our findings from Step 1. However, these results are not surprising to the extent that eligible cohort size (the denominator of the award concentration variable) is correlated with program size. As a program grows, the number of students eligible for the GRFP is likely to increase by more than the number of awarded students (the numerator), since award receipt remains a relatively rare event. The differential effect of the largest program size, Q4, (second largest program size, Q3) in reference to the smallest program size is associated with a 0.342 (0.393) standard deviation decrease in the GRFP award concentration.

In considering the set of results for the stratified samples (columns 2 and 3), the positive effects of faculty publication activity, GRE scores, a university high rank are moderated by highly ranked programs, while the negative results for the larger program sizes are

Table 3. Average marginal effects of logistic estimation of Equation (1), Step 1

Variables	(1) Full sample	(2) Rank 1	(3) Ranks 2 and 3
Leadership quality and composition			
Standardized average publications per faculty	**0.072***** (0.022)	**0.079***** (0.028)	0.075 (0.053)
Standardized percent of faculty with grants	−0.002 (0.016)	0.026 (0.024)	−0.018 (0.024)
Standardized percent female faculty	−0.011 (0.015)	0.009 (0.020)	0.004 (0.025)
Standardized percent of interdisciplinary faculty	**0.030**** (0.014)	**0.031*** (0.018)	0.018 (0.025)
Standardized percent non-Asian minority faculty	0.000 (0.014)	0.008 (0.021)	0.004 (0.022)
Peer quality and composition			
Standardized average GRE quantitative score	**0.067***** (0.019)	**0.085***** (0.029)	**0.070**** (0.030)
Standardized percent female students	−0.005 (0.019)	−0.022 (0.027)	0.021 (0.028)
Standardized percent of non-Asian Minority students	0.013 (0.015)	0.024 (0.024)	0.001 (0.022)
Standardized percent of students with academic plans	0.008 (0.015)	−0.014 (0.022)	**0.046*** (0.024)
Program support and traits			
Student workspace provided (binary)	0.022 (0.038)	−0.008 (0.045)	0.066 (0.071)
Student proposal support provided (binary)	0.012 (0.032)	0.041 (0.041)	−0.029 (0.053)
Above field average in student support programs (binary)	**0.083***** (0.031)	**0.110***** (0.039)	0.074 (0.052)
Standardized median time to degree	−0.008 (0.016)	−0.013 (0.023)	−0.016 (0.024)
Program size quartiles (referent: Q1, smallest)			
Program size, Q2	0.057 (0.052)	**0.193*** (0.102)	0.081 (0.071)
Program size, Q3	**0.096*** (0.052)	**0.213**** (0.096)	**0.137*** (0.075)
Program size, Q4 (largest)	**0.220***** (0.053)	**0.355***** (0.095)	**0.193**** (0.086)
University traits			
Public university (binary)	**−0.148***** (0.042)	**−0.178***** (0.048)	−0.088 (0.080)
University rank (referent: low)			
University rank: high	0.021 (0.057)	0.004 (0.104)	−0.034 (0.084)
University rank: mid	−0.003 (0.055)	−0.041 (0.105)	−0.065 (0.076)
Observations	1,028	601	406
Academic field fixed effects	Yes	Yes	Yes
Region controls included	Yes	Yes	Yes

Notes: Logistic regression results; average marginal effects presented; Standard errors in parentheses; $^{***}P < 0.01$, $^{**}P < 0.05$, $^{*}P < 0.1$.

Statistically significant results are bolded. Outcome: binary indicator if any awards received comparing effects for programs with only honorable mentions to programs with at least one award. All standardized covariates are normalized by academic division. Due to perfect prediction, certain narrow fields were omitted from the estimations: For the full sample (1) Computer Science, Mathematics, Nutrition, and Statistics and Probability; for rank 1 programs (2) Computer Science, Nutrition, Statistics and Probability, Entomology, and Forestry and Forest Science; and for ranks 2 and 3 programs (3) Mathematics, Entomology, Communication, Economics, Oceanography and Atmospheric Sciences, Operations Research and Systems Engineering, and Pharmacology and Toxicology.

moderated by the lower-ranked programs. The standardized measure for the percent of interdisciplinary activity is significant for the high-ranked sample, though the β coefficient is relatively small (0.053).

4.3 Post-specification tests

We run a series of post-specification tests to assess the fit of the various models. The results from post estimation analyses indicate that multicollinearity does not pose an issue (Variance Inflation Factor

Table 4. Coefficients of OLS estimation of Equation (2), Step 2

Variables	(1) Full sample	(2) Rank 1	(3) Ranks 2 and 3
Leadership quality and composition			
Standardized average publications per faculty	**0.151********	**0.130****	0.149
	(0.049)	(0.066)	(0.115)
Standardized percent of faculty with grants	0.012	0.077	−0.017
	(0.045)	(0.048)	(0.066)
Standardized percent female faculty	0.013	0.046	0.046
	(0.034)	(0.040)	(0.059)
Standardized percent of interdisciplinary faculty	0.039	**0.053***	0.026
	(0.027)	(0.031)	(0.065)
Standardized percent non-Asian minority faculty	−0.023	−0.047	0.008
	(0.033)	(0.051)	(0.041)
Peer quality and composition			
Standardized average GRE quantitative score	**0.159****	**0.308*****	0.049
	(0.075)	(0.107)	(0.070)
Standardized percent female students	−0.043	−0.002	−0.043
	(0.055)	(0.081)	(0.078)
Standardized percent of non-Asian minority students	0.047	**0.130*****	−0.040
	(0.032)	(0.046)	(0.032)
Standardized percent of students with academic plans	0.059	0.102	0.054
	(0.044)	(0.075)	(0.060)
Program support and traits			
Student workspace provided (binary)	0.046	−0.050	0.120
	(0.058)	(0.085)	(0.117)
Student proposal support provided (binary)	0.059	0.120	0.024
	(0.063)	(0.089)	(0.114)
Above field average in student support	0.050	0.076	−0.023
Programs (binary)	(0.049)	(0.063)	(0.084)
Standardized median time to degree	0.057	0.045	0.066
	(0.043)	(0.032)	(0.066)
Program size quartiles (referent: Q1, smallest)			
Program size, Q2	−0.129	0.502	−0.242
	(0.149)	(0.351)	(0.176)
Program size, Q3	**−0.393*****	0.082	**−0.361*****
	(0.121)	(0.245)	(0.127)
Program size, Q4 (largest)	**−0.342*****	0.192	**−0.449*****
	(0.107)	(0.198)	(0.135)
University traits			
Public university (binary)	**−0.186****	−0.125	−0.125
	(0.079)	(0.091)	(0.163)
University rank (referent: low)			
University rank: high	**0.203***	**0.532****	−0.052
	(0.114)	(0.198)	(0.135)
University rank: mid	0.037	0.187	−0.131
	(0.132)	(0.169)	(0.153)
Constant	0.560***	−0.108	0.798***
	(0.189)	(0.267)	(0.216)
Observations	1,033	611	422
R^2	0.217	0.309	0.213
Academic field fixed effects	Yes	Yes	Yes
Region controls included	Yes	Yes	Yes

Notes: OLS regression results, clustered standard errors in parentheses; ***$P < 0.01$, **$P < 0.05$, *$P < 0.1$; statistically significant results are bolded; Outcome: standardized ratio of awards to eligible student cohort where the numerator is the average annual award count and the denominator is the average first-year cohort size multiplied by two to account for first- and second-year students' eligibility to apply. All standardized variables are normalized by academic division.

(VIF) < 10). Moreover, for the binary model in Step 1 (Equation (1)), we also fit the model to a probit distribution. In addition, we run the model as a linear probability model (LPM) with an OLS distribution to assess the consistency and the efficiency of the results. For the OLS model, we clustered the standard errors by academic field. The results are robust across these two additional models to the primary results.

In addition, we test the functional form of the covariates. While we present the standardized values for the set of continuous measures, we also estimated the continuous regressors as: (1) baseline continuous/count measures; (2) location quotients relative to broad academic division base; and (3), binary, tercile, and quartile rankings of the baseline measures by division. For baseline measures, we estimate Step 2 with a count distribution (both using Poisson and Negative Binomial distributions) with the number of awardees as the outcome model. The results are robust across functional forms.

While the baseline measures offer the most direct measure, of critical note, the wide distribution across the continuous measures (e.g. average faculty publications vs. average GRE scores) makes interpretation of the coefficients less meaningful. Importantly, we present the results with the standardized notation not only to allow for comparison across regressors, but also to normalize the interpretation of the marginal effects.

In addition, we also standardize the continuous variables across the full sample by program rank (rather than academic division). Both alternatives produced similar findings to the standardized form by academic division. Moreover, for all models we both excluded and included academic field fixed effects; this reflects the 41 S&E fields from the NRC study.

Although we control for field-level fixed effects, this sample draws upon four broad academic divisions. Hence, we apply additional stratifications and estimate the primary models by academic division, respectively (Engineering, Life Sciences, Math and Physical Sciences, and Social and Behavioral Sciences). Most notably, the results for faculty publication, interdisciplinary faculty, and program size are most robust[17] for each division to the full sample, though the sizes of coefficients slightly vary. Finally, given the variation in funding demands across divisions, we also ran the model removing the Social Sciences divisions from the sample; the results are robust.[18]

As a final effort, we restrict the sample to programs with any award activity *only*. Effectively, we estimate the GRFP award concentration on the same set of variables as reported in Steps 1 and 2. We find that faculty publication activity, GRE scores, and public university are no longer significant; however, peer effects are driven by those with academic plans and longer duration of programs increase the GRFP award concentration. For highly ranked programs, the peer effects are robust; in addition, GRE scores are also a positive and significant indicator of increased GRFP award activity. Moreover, high-ranked programs are also moderating the effect of high university rank. Whereas, the lower-ranked programs moderate the negative effects for larger program size.[19]

5. Discussion

We rely on competitive R&D funding variation between academic programs with GRFP honorable mention and award-winning students to investigate the importance of graduate program and university level factors on research funding success. Moreover, this study redirects attention from senior scholars to emerging researchers—an often-overlooked population at an earlier point in their professional careers (Lane and Bertuzzi 2011; Callier and Polka 2015; Gould 2015; Woolston 2015). While we do not definitively identify the treated and control samples given the empirical challenge of effectively controlling for underlying student quality, our findings point to the importance of higher education institutional characteristics, support, and reputation as mediators of emerging researcher funding receipt. Moreover, academic institutional factors—that are external to the proposal—arguably serve as recognizable quality signals to NSF reviewers and offer knowledge spillovers that increase proposal quality.

Of note, the results from Step 1 report consistent positive effects of faculty publications on program award activity. We anticipate that increased faculty research activity may either spillover within the program, thus impacting the graduate student. Or, this research activity may spillover across the larger academic community sending a positive signal regarding reputation of the program and thus influencing the panel's perception of the proposal's potential. Given the nature of the data, we are unable to tease apart the specific mechanisms; however, we suspect that both factor into the funding decision. Distinguishing between these mechanisms across the set of significant outcomes offers a fruitful line for future research.

Additionally, the results indicate that larger program size positively influences whether a program can expect student applicants to receive an award. While this measure directly counts student personnel within a program—and indirectly those eligible to apply for a GRFP—program size also serves as a useful proxy for the scale of resources and human capital potential within the program. Thus, we interpret this as evidence that programs with greater resources positively influence student grant receipt. Corroborating this finding, we find that being a program in a public university is associated with a decrease in the probability of award receipt. US public institutions, in contrast to private research universities, on average tend to have less access to resources, which impacts research productivity (Aghion et al. 2010) and follow on funding (Lanahan et al. 2016).

In Step 2, we adjust the outcome variable to examine the effect of the set of regressors on award concentration. The set of robust results closely mirror our findings from Step 1. However, by adjusting the outcome variable from binary to a continuous measure, we find that program size *decreases* the concentration of awards. This finding is not altogether surprising considering that our outcome variable is a fraction representing the number of awards (numerator) per eligible cohort (denominator) in a given program. As a program increases in size, the denominator is likely to increase at a greater rate than the numerator due to the highly competitive nature of the GRFP.

Still, it is noteworthy that the full sample effect is moderated by lower-ranked programs. Thus in contrast to the smallest sized programs (Q1), larger, lower-ranked programs are more likely to experience a lower share of award winners out of eligible students. If we continue to view program size as an indication of available research resources, this implies that larger, lower-quality programs may face coordination costs that are detrimental to the rate of graduate student success. While larger programs increase the likelihood of having any awards, the larger, lower-ranked programs appear to exhibit greater inefficiencies in scaling the activity. These findings open a line of future research to examine the underlying mechanisms that account for this scaling cost.

Taken together, we have evidence that a series of leadership, peer, programmatic, and university characteristics are associated with grant funding success for emerging researchers. The significant results indicate there are institutional factors that mediate an applicant's chance at receiving an award given a baseline of quality. As graduate students embark on a career in research, the academic setting plays a pivotal role not only in shaping their training, but also in providing access to resources to launch their own research agendas.

Graduate students and graduate programs alike may benefit in considering the implications of these results. Prospective students would do well to consider programs with the most available resources in terms of human capital, support services, and financial support. Graduate programs may better support graduate students by supporting faculty publication efforts and hiring more interdisciplinary faculty. These efforts may boost the reputation of the program and bolster the student-proposal's 'broader impacts' statement.[20] Further, this study suggests that lower-ranked, larger departments may find it valuable to scale their program *down* and improve resources for fewer students.

5.1 Public policy implications

The findings of this study also have potential implications for NSF and its administration of the GRFP program. Made aware of what university and programmatic level factors are associated with award receipt, NSF may consider whether these factors play a role by directly bolstering proposal content (e.g. clear, thorough explication of the research agenda) and/or by playing upon implicit biases of review panel members. Reviewers may favor applicants that are advised by highly productive faculty or enrolled in well-financed private institutions over otherwise equally qualified applicants from less-regarded programs. This bias may be cause for concern, but may also be seen as a rationale heuristic given the abbreviated nature of the proposal itself and the nascent research experience of student PIs. Reputable departments with greater human and physical capital resources at their disposal may be more likely to ensure the research project is executed effectively.

Still, given the scale and scope of the GRFP, it is important to consider whether this external grant serves as a substitute or complement for internal funding. Given the constraints that many programs—especially within public institutions—face, the GRFP can serve as a pivotal funding source. If private institutions remain the disproportionate beneficiaries of GRFP funding, the gap in the ability of private and public institutions to support researchers is likely to widen (Aghion et al. 2010; Whalley and Hicks 2014).

The issue of disparities between groups also arises if NSF were to consider how current award conferment patterns fit within its larger GRFP mission. Reaffirming its commitment to diversifying the STEM (science, technology, engineering, and mathematics) workforce, the GRFP program has amended its eligibility policy with the explicit aim of identifying and encouraging a more diverse set of applicants to apply.[21] Targeted groups include women, underrepresented minorities, disabled persons, and veterans. To the extent that these students disproportionately place in public institutions, in middle-to-low ranking programs, and/or in programs with constrained student resources, current award activity evidence suggests these students may still face obstacles in securing a GRFP.

5.2 Future considerations

As with any research project, there are limitations. It should be noted that we are looking at a limited timeframe due to the collection period of the NRC data. Additionally, we do not know definitively at which point in their graduate training they received the GRFP recognition. More detailed data on the background of the applicant and access to the GRFP applicant data would allow for additional research on how the graduate program characteristics motivate students to apply for external research funding. While NSF data on the full applicant pool are restricted, future research could conduct follow-up surveys of faculty and emerging researchers to illuminate the underlying mechanisms that account for the reported variation.

Further, analysis at a more granular level could advance the discussion. At the applicant level, one could also assess receipt of the GRFP award on other outcomes such as research productivity as measured by publication activity and professional placement. If award winners, compared with honorable mentions experience more substantial professional gains as a result of this early stage funding, disparities between graduate programs at the award decision stage may persist, or even amplify. Future research should investigate the degree to which this prestige impacts the student's professional trajectory.

The student's *continued* success remains a salient line of research with important educational and economic implications. Certain academic environments promote a culture that values success in obtaining these grants as they recognize the value of external funding to the student's own financial stability, to the support of their lab, and for the prestige to their program. These graduate programs are providing a supportive environment to students that allow them to capitalize on their individual research ideas.

Anecdotal evidence suggests that the positive effects of GRFP receipt are felt strongest for winners located within laboratory environments, which are better accustomed to collaborative research projects, or in academic fields where mentorship is more prevalent. Put another way, the student may be at a research disadvantage if the external funding distances their access to mentorship during the graduate training. This may be more prominent in certain fields. For example, an economics GRFP winner may become 'siloed' within the program, which would isolate them away from faculty or other students and collaborative research opportunities. Contrary to the aims of the program, this could reduce the student's research output. Though this article does not address the impacts of GRFP receipt, these potential implications are important to consider. Further assessment is necessary to examine how the award affects the future success of both the emerging researchers and their programs.

Acknowledgements

We would like to thank Jonathan Eyer, Daniel Armanios, and anonymous reviewers for comments on earlier versions of the article. Moreover, we thank Emily Plews, Georgiy Sichinava, Soro Soukpafolo, and Zachary Trudo for their research assistance on the larger research project. This article benefitted from discussions with seminar participants at the 2015 Atlanta Conference on Science and Innovation Policy, the 2015 Price Junior Faculty Research Meeting, and the 2016 Association of Public Policy and Management Annual Meeting. All errors are our own.

Funding

This work was supported in part by the National Science Foundation Science of Science and Innovation Policy Program [1548288]. NSF had no role in the design, analysis, or writing of this research.

Notes

1. http://www.sciencemag.org/news/2015/12/updated-budget-agreement-boosts-us-science
2. http://www.aaas.org/fy16budget/federal-rd-fy-2016-budget-overview#rd
3. Higher education institutions are the context of this study. These institutions are composed of nested levels with the program (analogous with department) as the most

granular, followed by broad field divisions (e.g. the Life Sciences), and finally the university level.
4. http://www.nsf.gov/statistics/2016/nsb20161/#/report/chapter-2/graduate-education-enrollment-and-degrees-in-the-united-states
5. Although the reputation of the higher education institution may vary in contrast to the quality and rank of the academic programs nested within the institution, evidence from the NRC's survey of doctorate programs indicates a positive correlation between these levels of the academic institution (0.356). This correlation is driven most significantly by the high prevalence of top ranked programs within top ranked universities.
6. GRFP Program Solicitation, NSF 15-597, http://www.nsf.gov/pubs/2015/nsf15597/nsf15597.htm
7. Although students can apply for the GRFP concurrent with graduate school applications, their status as prospective graduate students may affect their likelihood of winning an award. For the purposes of this analysis, we are centrally concerned with organizational factors from the *graduate* institution on funding receipt. As such, we relied on the GRFP award database to identify students currently enrolled in graduate institutions (i.e. those in either their first or second year of graduate school). We based this on GRFP proposals where the 'Proposed' and 'Current' institutions match. See Appendix, section *Step 1: GRFP Data* for more detail.
8. GRFP Program Solicitation, NSF 15-597, http://www.nsf.gov/pubs/2015/nsf15597/nsf15597.htm
9. The NRC built a taxonomy of academic fields for this study designated most prominently by the following: '[the field] must have produced at least 500 Ph.D.s over the most recent 5 years and be offered by programs that had produced 5 or more Ph.D.s in the last 5 years in at least 25 universities' (49: p. 19). Of note, the fields are composed of relevant subfields that correspond most directly with academic departments. See Appendix, section *Building Program-Level Dataset* for more detail.
10. The number of awards increased from 1,248 in 2009 to 2,051 in 2010.
11. The ratio of honorable mentions to awards decreased from an average of 1.63 (2000–9) to 0.89 (2010–14).
12. As a sensitivity measure, we estimate Equation (1) with an alternate bottom threshold of at least two awards. The results are presented in Appendix Table A.5 and are consistent and at times stronger. However, this restriction decreases the sample by 5–30% depending on the control group definition so we report for the larger more conservative sample as our primary results.
13. The share of programs with only awards is 12% (127).
14. Graduate student support programs include: student orientation, international student orientation, language support, writing support, statistics support, prizes for teaching or research, proposal support, on-campus graduate research conference, academic integrity training, graduate student association, staff and graduate student association, financial support of graduate student association, academic grievance support, dispute resolution, regular graduate program meeting, annual review, teacher training, and travel support. Refer to the Appendix for more detail.
15. We use university rankings from Barron's Profiles of American Colleges Online Edition (http://www.BarronsPAC.com)—document number: 4-0650 PAC pt4_1.
16. Refer to Appendix Table A.2.
17. Of note, the direction of the coefficient for interdisciplinary faculty is negative for the Math and Physical Sciences stratification.
18. The results are reported in Table A.6 in the Appendix.
19. The results for each of these post specification tests are available upon request.
20. Broader impacts are given equal weight to a proposal's 'intellectual merit' statement.https://www.nsfgrfp.org/applicants/application_components/merit_review_criteria
21. Effective the 2017 competition, NSF has restricted graduate students to only one application—submitted either in the first *or* second year of graduate school—with the goal of recruiting more diverse undergraduates (see https://www.nsf.gov/pubs/2016/nsf16050/nsf16050.pdf).
22. NSF reports data on six fields: Name of applicant, Email Address, Baccalaureate Institution, Field of Study, Proposed Graduate Institution, and Current Institution, though not all fields appear for each year of awardee and honorable mention data available. Email Address and Current Institution fields are most complete after 2003.
23. NSF GRFP proposal data available at: https://www.fastlane.nsf.gov/grfp/AwardeeList.do?methodloadAwardeeList.
24. This sample represents 99.22% of GRFP awardees and honorable mentions. Reporting for honorable mentions is 100% from 1994 to 2004. For more recent years (2005–14), the average reporting ratio is 97.69%. The overall reporting ratio for honorable mentions is 98.90%. This was derived from the following source: https://www.fastlane.nsf.gov/grfp/AwardeeList.do?methodloadAwardeeList. In all, 457 observations were dropped due to missing data on one of the following fields: Field of Study, Proposed Graduate Institution, or Current Institution.
25. http://sites.nationalacademies.org/PGA/Resdoc/PGA_044522
26. Data on the population of US academic institutions are publicly available: https://nces.ed.gov/ipeds/datacenter/.
27. http://sites.nationalacademies.org/PGA/Resdoc/PGA_044522
28. The results for these models are available upon request.
29. The authors employed a number of procedures to match on string name including removing spaces, renaming common terms (e.g. 'university' 'uni'; 'college' 'col' 'coll'; 'inst' 'institute'; 'state' 'st'), and removing articles (e.g. 'and', 'of', 'at', 'for', ect.).
30. This reflects the number of Current Institutions based on a unique string sequence. There were 974 unique string sequences for the Proposed Graduate Institutions.
31. We suspect that those who listed a foreign institution

were applying for the GRFP award while applying for a graduate program at a US-based institution. At the time of submitting the grant application; however, they were located at another institution. Given our emphasis on examining the effect of graduate programs on research promise, we removed this sample.

32. For example, the Public Health program in the Biological and Sciences field was listed at both San Diego State University and the University of California San Diego (http://publichealth.sdsu.edu); the Biomedical Engineering and Bioengineering program in the Engineering field was listed at both the Georgia Institute of Technology and Emory University http://www.bme.gatech.edu; and the Civil and Environmental Engineering program in the Engineering field was listed at the University of Alabama Birmingham and the University of Alabama in Huntsville (www.eng.uab.edu/cee).
33. http://sites.nationalacademies.org/pga/resdoc/
34. NRC sixty-two fields include: Aerospace Engineering; Agricultural and Research Economics; Animal Sciences; Anthropology; Applied Mathematics; Astrophysics and Astronomy; Biochemistry, Biophysics, and Structural Biology; Biology/Integrated Biology/Integrated Biomedical Sciences; Biomedical Engineering and Bioengineering; Cell and Developmental Biology; Chemical Engineering; Chemistry; Civil and Environmental Engineering; Communication; Computer Engineering; Computer Sciences; Earth Sciences; Ecology and Evolutionary Biology; Economics; Electrical and Computer Engineering; Engineering Science and Materials; Entomology; Food Science; Forestry and Forest Sciences; Genetics and Genomics; Geography; History; Immunology and Infectious Disease; Kinesiology; Linguistics; Materials Science and Engineering; Mathematics; Mechanical Engineering; Microbiology; Neuroscience and Neurobiology; Nursing; Nutrition; Oceanography, Atmospheric Sciences and Meteorology; Operations Research, Systems Engineering and Industrial Engineering; Pharmacology, Toxicology and Environmental Health; Physics; Physiology; Plant Sciences; Political Science; Psychology; Public Affairs, Public Policy and Public Administration; Public Health; Sociology; Statistics and Probability; American Studies; Classics; Comparative Literature; English Language and Literature; French and Francophone Language and Literature; German Language and Literature; History of Art, Architecture and Archaeology; Languages, Societies and Cultures; Music (except performance); Philosophy; Religion; Spanish and Portuguese Language and Literature; and Theatre and Performance Studies.
35. NRC five fields include: Engineering, Life Sciences, Physical and Mathematical Sciences, Social and Behavioral Sciences, and Arts and Humanities.
36. Arts and Humanities broad field includes: American Studies; Classics; Comparative Literature; English Language and Literature; French and Francophone Language and Literature; German Language and Literature; History; History of Art, Architecture and Archaeology; Languages, Societies and Cultures; Music (except performance); Philosophy; Religion; Spanish and Portuguese Language and Literature; and Theatre and Performance Studies.
37. Health-related fields: Nursing, and Kinesiology.
38. Source: http://sites.nationalacademies.org/PGA/Resdoc/PGA_044478
39. A total of 538 observations reported 'N/A' for program.
40. Data on the population of US academic institutions are publicly available: https://nces.ed.gov/ipeds/datacenter/.
41. http://sites.nationalacademies.org/PGA/Resdoc/PGA_044522
42. When merging the NSF GRFP data to the NRC data, 62.1% of the NRC sample had some NSF GRFP activity over the twenty-year timeframe; thus indicating that 37.9% of the programs surveyed by the NRC had no GRFP activity over the twenty-year timeframe.
43. There were 612 duplicates. Based on the nature of the covariate, we either summed or took the average of the duplicate observations. As illustrative examples of the duplicates, Rice University had two narrow programs listed as 'Civil and Environmental Engineering' to reflect the Civil Engineering and Environmental Engineering programs at the institution; the University of Southern California had four narrow programs listed as 'Public Health' to reflect the Biometry, Epidemiology, Occupational Science, and Preventative Medicine (Health Behavior) programs at the institution; and Yale University has two narrow programs listed as 'Immunology and Infectious Disease' to reflect the Experimental Pathology and Immunobiology programs at the institution. The taxonomy of fields presented by Board on Higher Education and Workforce at the National Academies represents a commendable effort to classify the extensive range of academic programs. To examine broad trends across programs, we rely on this effort and aggregate the duplicates where multiple programs comprise a narrow field.

References

Agarwal, R., and Ohyama, A. (2013) 'Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists', *Management Science*, 59/4: 950–70.

Aggarwal, V., Hsu, D., and Wu, A. (2015). 'R&D Production Team Organization and Firm-Level Innovation'. Working Paper.

Aghion, P. et al. (2010) 'The Governance and Performance of Universities: Evidence from Europe and the US', *Economic Policy*, 25/61: 7–59.

Aldrich, H., and Fiol, C. (1994) 'Fools Rush In? The Institutional Context of Industry Creation', *Academy of Management Review*, 19/4: 645–70.

Amar, A., Hentrich, C., and Hlupic, V. (2009) 'To Be a Better Leader, Give Up Authority', *Harvard Business Review*, December.

Amin, A., and Cohendet, P. (2000) 'Organisational Learning and Governance through Embedded Practices', *Journal of Management and Governance*, 4: 93–116.

Arora, A., and Gambardella, A. (2005) 'The Impact of NSF Support for Basic Research in Economics', *Annales d'Economie et de Statistique*, July/December: 91–117.

Autio, E. et al. (2014) 'Entrepreneurial Innovation: The Importance of Context', *Research Policy*, 43/7: 1097–108.

Azoulay, P., Ding, W. and Stuart, T. (2007) 'The Determinants of Faculty Patenting Behavior: Demographics or Opportunities?', *Journal of Economic Behavior & Organization*, 63/4: 599–623.

——, Graff Zivin, J., and Manso, G. (2011) 'Incentives and Creativity: Evidence from the Academic Life Sciences', *The RAND Journal of Economics*, 42: 527–54.

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bartolone, J. et al. (2014) *Evaluation of the National Science Foundation's Graduate Research Fellowship Program – Final Report*. Chicago, IL: NORC at the University of Chicago.

Bercovitz, J., and Feldman, M. (2008) 'Academic Entrepreneurs: Organizational Change at the Individual Level', *Organization Science*, 19/1: 69–89.

Bozeman, B., and Corley, E. (2004) 'Scientists' Collaboration Strategies: Implications for Scientific and Technical Human Capital', *Research Policy*, 33/4: 599–616.

Brooks, R. (2005) 'Measuring University Quality', *The Review of Higher Education*, 29/1: 1–21.

Callier, V., and Polka, J. (2015) 'Fellowships are the Future', *Nature*, 528/7580: 155–6.

Cetina, K. (2009) *Epistemic Cultures: How the Sciences make Knowledge*. Cambridge, MA: Harvard University Press.

Chapman, G., and McCauley, C. (1993) 'Early Career Achievements of National Science Foundation (NSF) Graduate Applicants: Looking for Pygmalion and Galatea Effects on NSF Winners', *Journal of Applied Psychology*, 78/5: 8–15.

Conti, A., and Liu, C. (2015) 'Bringing the Lab Back in: Personnel Composition and Scientific Output at the MIT Department of Biology', *Research Policy*, 44/9: 1633–44.

de Valero, Y. (2001) 'Departmental Factors Affecting Time-to-Degree and Completion Rates of Doctoral Students at One Land-Grant Research Institution', *Journal of Higher Education*, 72:341–67.

Dill, D. (2006) 'Convergence and Diversity: The Role and Influence of University Rankings'. In Keynote Address presented at the Consortium of Higher Education Researchers (CHER) 19th Annual Research Conference, September (Vol. 9).

—— and Beerkens, M. (2013) 'Designing the Framework Conditions for Assuring Academic Standards: Lessons Learned about Professional, Market, and Government Regulation of Academic Quality', *Higher Education*, 65/3: 341–57.

DiMaggio, P., and Powell, W. (1983) 'The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields', *American Sociology Review*, 48/2: 147–60.

Duflo, E., and Saez, E. (2000) 'Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues Choices'. Working Paper 00-07 MIT Department of Economics, Cambridge, MA.

Ellison, G., and Fudenberg, D. (1993) 'Rules of Thumb for Social Learning', *Journal of Political Economy*, 101/4: 612–43.

Etzkowitz, H. (2003) 'Research Groups as "quasi-Firms": The Invention of the Entrepreneurial University', *Research Policy*, 32/1: 109–21.

Friedman, R. S., and Friedman, R. C. (1982) *The Role of University Organized Research Units in Academic Science*. University Park: Pennsylvania State University Institute for Policy Research and Evaluation.

Gardner, S. (2009) 'Conceptualizing Success in Doctoral Education: Perspectives of Faculty in Seven Disciplines', *The Review of Higher Education*, 32/3: 383–406.

Goldfarb, B. (2008) 'The Effect of Government Contracting on Academic Research: Does the Source of Funding Affect Scientific Output?', *Research Policy*, 37: 41–58.

Gould, J. (2015) 'How to Build a Better PhD', *Nature*, 528: 22–26.

Hegde, D. (2005) 'Public and Private Universities: Unequal Sources of Regional Innovation?', *Economic Development Quarterly*, 19/4: 373–86.

Heinze, T. (2008) 'How to Sponsor Ground-Breaking Research: A Comparison of Funding Schemes', *Science and Public Policy*, 35/5: 302–18.

Hicks, D. (2009) 'Evolving Regimes of Multi-university Research Evaluation', *Higher Education*, 57/4: 393–404.

Jacobs, J., and Frickel, S. (2009) 'Interdisciplinarity: A Critical Assessment', *Annual Review of Sociology*, 35: 43–65.

—— and Lefgren, L. (2011a) 'The Impact of Research Grant Funding on Scientific Productivity', *Journal of Public Economics*, 95: 1168–77.

—— and —— (2011b) 'The Impact of NIH Postdoctoral Training Grants on Scientific Productivity', *Research Policy*, 40: 864–74.

Jones, B., Wuchty, S., and Uzzi, B. (2008) 'Multi-university Research Teams: Shifting Impact, Geography, and Stratification in Science', *Science*, 322: 1259–62.

Lanahan, L., Graddy-Reed, A., and Feldman, M. (2016) 'The Domino Effects of Federal Research Funding', *PLOS One*, 11/6: e0157325.

Lane, J., and Bertuzzi, L. (2011) 'Measuring the Results of Science Investments', *Science*, 331/6018: 678–80.

—— et al. (2015) 'New Linked Data on Research Investments: Scientific Workforce, Productivity, and Public Value', *Research Policy*, 44: 1659–71.

Lattuca, L. (2001) *Creating Interdisciplinarity: Interdisciplinary Research and Teaching among College and University Faculty*. Nashville, TN: Vanderbilt University Press.

Le, V., and Bartolone, J. (2015) *Employment and Professional Productivity of Graduate Research Fellowship Program Fellows (1994–2011)*. Chicago, IL: NORC at the University of Chicago.

Nature (2015) 'Make the most of PhDs', *Nature*, 528: 7.

Nerad, M., and Cerny, J. (1993) 'From Facts to Action: Expanding the Graduate Division's Educational Role', *New Directions for Institutional Research*, 1993/80: 27–39.

Organization for Economic Co-operation and Development (2012) *Education at a Glance 2012*. Paris, France: OECD Publishing.

Ostriker, J. et al. (2011) *A Data-Based Assessment of Research-Doctorate Programs in the United States*. Washington, DC: National Academies Press

Pinheiro, D., Melkers, J., and Youtie, J. (2014) 'Learning to Play the Game: Student Publishing as an Indicator of Future Scholarly Success', *Technological Forecasting and Social Change*, 81: 56–66.

Schmitt, S. (2013) 'Assessment and Rankings Efforts: The Effect on Institutional and Program-Level Change', Proquest Dissertations and Theses Database.

Shamir, B., House, R., and Arthur, M. (1993) 'The Motivational Effects of Charismatic Leadership: A Self-Concept Based Theory', *Organization Science*, 4/4: 577–94.

Singh, J., and Fleming, L. (2010) 'Lone Inventors as Sources of Breakthroughs: Myth or Reality?', *Management Science*, 56/1: 41–56.

Sorensen, A. (2002) *Social Learning in the Demand for Employer Sponsored Health Insurance*. Baltimore, MD: Johns Hopkins University Press.

Stephan, P. (2012) *How Economics Shapes Science* (Vol. 1). Cambridge, MA: Harvard University Press.

Stinchcombe, A. (1965) 'Organizations and Social Structure', *Handbook of Organizations*, 44/2: 142–93.

Whalley, A., and Hicks, J. (2014) 'Spending Wisely? How Resources Affect Knowledge Production in Universities', *Economic Inquiry*, 52/1: 35–55.

Williams, K., and O'Reilly, C. (1998) 'Demography and Diversity in Organizations: A Review of 40 Years of Research', *Research in Organizational Behavior*, 20: 77–140.

Woolston, C. (2015) 'Uncertain Futures', *Nature*, 526: 597–600.

Wuchty, S., Jones, B., and Uzzi, B. (2007) 'The Increasing Dominance of Teams in Production of Knowledge', *Science*, 316/5827: 1036–9.

Zucker, L., and Darby, M. (1996) 'Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry', *Proceedings of the National Academy of Sciences*, 93/23: 12709–16.

Appendix

Building the program-level dataset

We use data from two separate databases: the NRC's data on Research Doctorate Programs and the National Science Foundation's GRFP grant database. The NRC database comprises a cross section of program observations, while the GRFP database comprises annual proposal observations.

The GRFP data is a time series dataset that is structured at the proposal-level and includes institutional affiliation and field of study for the population of awardees and honorable mentions, respectively.[22] Data on GRFP award and honorable mention activity were accessed publicly online.[23] At the time of building this dataset, data on awardees was available from 1952 to 2014, and data on honorable mentions were available from 1994 to 2014. We built a database based on availability of data for the population of awardees and honorable mentions, therefore delimiting the timeframe from 1994 to 2014. This includes 58,218 unique observations.[24] While this timeframe exceeds the four-year period used in this analysis, we built the database based on the full sample of data and then restricted the sample from 2005 to 2008 ex post for the empirical analysis.

The NRC research doctorate study reports program-level data for 5,004 program at 212 universities. The survey reports program-level statistics for activity between 2000 and 2006; the results from the survey were made publicly available in 2010. The NRC study provides a representative database of doctoral-research programs (Ostriker et al. 2011), which is revered as the most comprehensive data source on US graduate programs. NRC programs with any GRFP activity define the sample graduate programs examined for this analysis.

The academic program is the level of analysis for this article, indexed by i for the academic field and n for the university. The 41 fields in the full sample comprise four broad academic divisions. Although this dataset presents comprehensive granular data on academic organization, it is important to note that the NRC fields are composed of corresponding subfields that more directly correspond with an academic department.[25] Table A.2 details the crosswalk by academic division and field.

To merge the two datasets, we first needed to create a unique university-field crosswalk. We referred to the NCES IPEDS[26] and the NRC taxonomy of academic fields and subfields[27] to assign a unique institutional identification and program identification, respectively. We used these two identifications to uniquely define the university-field in each dataset before merging the two datasets together.

Below, we outline the three main steps to build this dataset. First, we identified the university IPEDS ID (Step 1) and academic field (Step 2) for the observations in each dataset respectively. For the former, we specify the parameters for identifying the graduate institution for the GRFP database as well. Then, we merged the two datasets on the unique numeric university-field crosswalk (Step 3). We coded and documented the entire match procedure in STATA.

Step 1: Assign University IPEDS ID

GRFP Data

Before assigning an IPEDS ID to the university, we relied on information from the fields 'Current Institution' and 'Proposed Graduate Institution' to define the university for the GRFP data. For the primary analysis, we included only those observations where the applicant's Proposed Graduate Institution matched the Current Institution. With an emphasis on examining the effect of graduate programs on emerging researcher activity, we are interested in identifying the subset of individuals who applied for the GRFP program while enrolled in his/her graduate program. To reiterate, emerging researchers are eligible to apply for the GRFP prior to completing twelve academic months of graduate study. Thus, individuals completing their undergraduate degrees or master's degree en route to applying to a PhD program are eligible. In these cases, the Current Institution and Proposed Graduate Institution may not match given that they may apply for the GRFP while they simultaneously apply to a graduate program at another institution. By limiting the population to those whose Proposed Graduate Institution match the Current Institution, we assume to identify those currently in a graduate program or at least exposed to the graduate program.

As additional sensitivity checks, we also identified the university based on the Current Institution and Proposed Graduate Institution, respectively. In these two cases, we assume that the Current Institution or the Proposed Graduate Institution is the individual's graduate institution. We recognize that these are weaker assumptions for two reasons: (1) the Current Institution may reflect the individual's undergraduate or master's institution—which may be different from his/her PhD institution; or (2) the individual may not end up enrolling at the listed Proposed Graduate Institution. The results were generally robust to the main models.[28] This suggests that GRFP honorable mentions and awardees are typically acknowledged once enrolled their graduate program. For the primary analysis we report where the Current Institution matches the Proposed Graduate Institution as it more accurately captures the population of GRFP applicants enrolled in and exposed to the graduate program when they receive GRFP recognition.

To match the IPEDS ID with the GRFP university string variable, we wrote a code to define the unique core sequence of strings for each institution. Articles were dropped and common terms were recoded.[29] This exercise defined 895 unique universities.[30] We employed the same method for cleaning the population of universities listed in the IPEDS database. As a result, we defined 15,679 unique academic organizations in the IPEDS database. We merged the datasets based on the unique university string variable. Approximately 72 percent GRFP observations were directly matched. For the remaining observations, we hand-matched 8,258 observations (27 percent) based on common sequences of strings. Among those that we were unable to assign an IPEDS ID 270 observations were foreign institutions[31] and 45 observations listed institutions with no IPEDS id. In sum, we matched IPEDS ID for 97.9 percent of the total GRFP sample.

NRC Data

We employed the same coding technique as noted above in terms of cleaning up the string variable to match universities in the NRC sample to IPEDS. For the first automated round of matching we directly matched 3,918 observations (78.3% of the sample). We hand-matched an additional 985 observations (19.7%), which yielded a 98.0% match rate. It is worth noting that 31 of the NRC programs were affiliated with more than one university.[32] We dropped these from the sample given their unique multi-university affiliation.

Step 2: Assign Field ID

NRC Data

We based the field ID classification directly on the NRC program listing, as specified by the broad field and narrow field classifications,[33] corresponding to division and field, respectively, in the article. This original dataset includes sixty-two narrow fields[34] from five broad fields.[35] Given the S&E scope of the NSF GRFP program, we removed fourteen fields from the NRC dataset that comprise the Arts and Humanities broad field,[36] and two fields with a health-related focus.[37] This yielded a total of forty-six S&E fields.

GRFP Data

To match the GRFP graduate program with the list of forty-six NRC fields, we defined a unique sequence of common strings based on the NRC's taxonomy of fields and their subfields.[38] For example, for the NRC field 'aerospace engineering' we identified a combination of the following strings as a common match: 'aeronautical vehicles', 'space vehicles', 'systems engineering and multidisciplinary design optimization', 'aerodynamics and fluid mechanics', 'astrodynamics', 'structures and materials', 'propulsion and power', 'navigation, guidance, control and dynamics', and 'multi-vehicle systems and air traffic control', As another example, for the NRC field 'genetics and genomics' we identified a combination of the following strings as a common match: 'computational biology', 'genetics', 'genomics', and 'molecular genetics'. As a result of this string-based approach, we matched NRC fields to GRFP graduate programs for 57,598 observations yielding a 98.9% match rate.[39]

Step 3: Creating a university-field crosswalk

To merge the GRFP and NRC data, we created a university-field numeric identification based on the publicly available NCES IPEDS[40] and NRC S&E fields.[41] We then collapsed the GRFP data to the program level to compute annual aggregate counts of awardees and honorable mentions. In preparation for the merge, we dropped all observations without both a university and field for each dataset respectively. This resulted in 3,951 programs.

When merging the GRFP data to the NRC data, 49.2% of the NRC sample had some NSF GRFP activity over four-year timeframe, 2005–8 leaving 50.8% with no NSF GRFP activity.[42] As a side note, roughly 44% of the GRFP sample did not match to the NRC sample. We attribute this nonmatch based on the restrictions of the sample for the NRC dataset. Notably among those that did not match, they represent a large distribution of universities (310); this in fact exceeds the sample size of the NRC survey. In addition, these unmatched GRFP data overrepresented six narrow fields: Public Affairs, Public Policy and Public Administration; Operations Research, Systems Engineering and Industrial Engineering; Mathematics; Engineering Science and Materials (not elsewhere classified); Communication; Biochemistry, Biophysics, and Structural Biology. Taken together, this suggests that the true distributions of US graduate programs for these six fields are larger than the distributions sampled by the NRC study. Importantly, the NRC study includes 5,004 programs, yet this distribution only represents 212 universities. On average then, 23 programs were surveyed from each institution, yet 62 programs in total were reviewed. Based on our assessment of the missing data, the GRFP data that did not merge to the NRC are programs that were not surveyed in the NRC study. We are unable to include this set in the analysis given that we do not have program-level data for the full population of US graduate research programs. While the NRC study did not survey every program in the USA, it is recognized as the most comprehensive and detailed survey of US research graduate programs.

After the merge, we assessed the data and discovered a handful of NRC program duplicates.[43] For those with more than one unique university-field, we computed an aggregate mean for the program observation and removed the duplicate values.

The sample of NRC programs with *any* GRFP award and/or honorable mention activity by students from 2005 to 2008 defines the dataset. Based on these restrictions, our full sample includes 1,033 graduate programs from 142 universities. This sample of programs secured 1,984 awards and 3,642 honorable mentions, for a total of 5,626 GRFP accolades from 2005 to 2008.

Appendix Tables

Table A1. NRC program sample frequency

	Program	Frequency	Percent
1	Aerospace Engineering	17	1.65
2	Animal Sciences	23	2.23
3	Anthropology	23	2.23
4	Applied Mathematics	2	0.19
5	Astrophysics and Astronomy	24	2.32
6	Biochemistry, Biophysics, and Structural Biology	38	3.68
7	Biology/Integrated Biology/Integrated Biomedical Sciences	31	3
8	Biomedical Engineering and Bioengineering	48	4.65
9	Cell and Developmental Biology	34	3.29
10	Chemical Engineering	32	3.1
11	Chemistry	54	5.23
12	Civil and Environmental Engineering	50	4.84
13	Communication	5	0.48
14	Computer Sciences	1	0.1
15	Earth Sciences	52	5.03
16	Ecology and Evolutionary Biology	50	4.84
17	Economics	18	1.74
18	Electrical and Computer Engineering	63	6.1
19	Entomology	8	0.77
20	Forestry and Forest Sciences	5	0.48
21	Genetics and Genomics	24	2.32
22	Geography	15	1.45
23	Immunology and Infectious Disease	16	1.55
24	Linguistics	22	2.13
25	Materials Science and Engineering	36	3.48
26	Mathematics	1	0.1
27	Mechanical Engineering	60	5.81
28	Microbiology	26	2.52
29	Neuroscience and Neurobiology	35	3.39
30	Nutrition	2	0.19
31	Oceanography, Atmospheric Sciences, and Meteorology	8	0.77
32	Operations Research, Systems Engineering, and Industrial Engineering	12	1.16
33	Pharmacology, Toxicology and Environment Health	15	1.45
34	Physics	15	1.45
35	Physiology	12	1.16
36	Plant Sciences	14	1.36
37	Political Science	44	4.26
38	Psychology	67	6.49
39	Public Affairs, Public Policy and Public Administration	3	0.29
40	Sociology	27	2.61
41	Statistics and Probability	1	0.1
	Total	1,033	100

Table A2. Crosswalk between academic division and academic field

Academic division	Academic field
Engineering	Aerospace Engineering; Biomedical Engineering and Bioengineering; Chemical Engineering, Civil and Environmental Engineering, Electrical and Computer Engineering; Materials Science and Engineering, Mechanical Engineering, Operations Research, Systems Engineering, and Industrial Engineering
Life Sciences	Animal Sciences; Biochemistry, Biophysics, and Structural Biology; Biology/Integrated Biology/Integrated Biomedical Sciences; Cell and Developmental Biology; Ecology and Evolutionary Biology; Entomology; Forestry and Forest Sciences; Genetics and Genomics; Immunology and Infectious Disease; Microbiology; Neuroscience and Neurobiology; Nutrition; Pharmacology, Toxicology, and Environmental Health; Physiology; Plant Sciences
Mathematics and Physical Sciences	Applied Mathematics; Astrophysics and Astronomy; Chemistry; Computer Sciences; Earth Sciences; Mathematics; Oceanography, Atmospheric Sciences, and Meteorology; Physics; Statistics and Probability
Social and Behavioral Sciences	Anthropology; Communications; Economics; Geography; Linguistics; Political Science; Psychology; Public Affairs, Public Policy and Public Administration; Sociology

Note: The crosswalk follows NRC's Program Classification (http://sites.nationalacademies.org/PGA/Resdoc/PGA_044522).

Table A3. Description of variables

Variable	Definition
Outcome variables	
Any award activity in 2005–8 (Step 1)	Binary variable equal to 1 if the program has any doctoral students receive a GRFP award between 2005 and 2008.
Share of awards to eligible students in 2005–8 (Step 2)	Ratio of the following: average number of annual GRFP awards received per year by doctoral students within the program between 2005 and 2008 (numerator) over double the size of the average first-year cohort as measured from 2002 to 2006 (denominator).
Leadership quality	
Average number of publications (2000–6) per allocated faculty, 2006	Average number of publications published by allocated faculty members in a program between 2000 and 2006 as of data collection in 2006 by Thomson Reuters. See page 241 in Ostriker et al. (2011) for more information.
Percent of faculty with grants, 2006	Number of faculty with extramural grant or contract support divided by the total number of faculty in the program in 2006.
Leadership composition	
Percent female faculty, 2006	Number of female core and new faculty members as of 2006 as a percent of total core and new faculty members in the program (not including allocated faculty).
Percent of interdisciplinary faculty, 2006	Number of faculty 'associated' with the program and at least one other program divided by the total number of faculty comprising that program including associated, core, and new faculty members in 2006.
Percent non-Asian minority faculty, 2006	Number of non-Hispanic Blacks, Hispanics, and American Indians or Alaskan Natives as of 2006 as a percent of total core and new faculty members including non-Hispanic Whites and Asians or Pacific Islanders in the program (not including allocated faculty). Does not include faculty with unknown race/ethnicities, non-US citizens, and nonpermanent residents.
Peer quality	
Average GRE quantitative scores, 2004–6	Weighted average quantitative GRE score for the program, calculated by multiplying the number of individuals reporting scores each year (2004, 2005, and 2006, separately) by the reported average GRE score for that year, summing these three quantities together, and dividing by the total sum of individuals reporting scores between 2004 and 2006 in that program.
Peer composition	
Percent female students, 2005	Number of female students enrolled in the program as of Fall 2005 as a percent of the total number of graduate students in the program.
Percent of non-Asian minority students, 2005	Number of non-Hispanic Blacks, Hispanics, and American Indians or Alaskan Natives students enrolled in the program as of Fall 2005 as a percent of the total number of domestic graduate students in the program with known race/ethnicities. Does not include faculty with unknown race/ethnicities, non-US citizens, and nonpermanent residents.
Percent of students with academic plans, 2001–5	Percent of doctoral students with definite plans for an academic position, calculated via a crosswalk with the NSF Doctorate Records File using data between 2001 and 2005. The percentage represents the number of individuals who signed or were negotiating a contract for a position at an educational institution in that field divided by the total number of survey respondents in that field between 2001 and 2005. See page 245 in Ostriker et al. (2011) for more information.
Program support	
Student workspace provided	Binary variable equal to 1 if the program provides students with workspace.
Proposal support available	Binary variable equal to 1 if the program or university provides students with proposal writing assistance or training.
Ratio of student support programs to field average indicator	Binary variable equal to 1 if the ratio of number of student activities offered compared to average number offered within academic field is above average, or greater than 1. This is based off of the eighteen potential student activities listed in the NRC, which includes: student orientation, international student orientation, language support, writing support, statistics support, prizes for teaching or research, proposal support, on-campus graduate research conference, academic integrity training, graduate student association, staff and graduate student association, financial support of graduate student association, academic grievance support, dispute resolution, regular graduate program meeting, annual review, teacher training, travel support
Median time to degree, 2004–6	Median time to degree for full-time and part-time students in the program measured in years and averaged over the years 2004–6.
Program size quartile ranking, Fall 2005	Categorical variable equal to the quartile ranking of the program based on the number of students enrolled as of Fall 2005; 1 is smallest while 4 is largest.
University traits	
Region	Categorical variable of the region of the university organized by state. *Northeast*: Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, Pennsylvania, New Jersey. *Midwest*: Wisconsin, Michigan, Illinois, Indiana, Ohio, Missouri, North Dakota,

(Continued)

Table A3. (Continued)

Variable	Definition
	South Dakota, Nebraska, Kansas, Minnesota, Iowa. *South Atlantic*: Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida. *South Central*: Kentucky, Tennessee, Mississippi, Alabama, Oklahoma, Texas, Arkansas, Louisiana. *West*: Idaho, Montana, Wyoming, Nevada, Utah, Colorado, Arizona, New Mexico, Alaska, Washington, Oregon, California, Hawaii.
Public university	Binary variable equal to 1 if the university is a public institution rather than privately controlled.
University rank	Categorical variable of the Barron's university rankings grouped into three tiers of: most and highly competitive, very competitive, or competitive and less competitive.

Notes: All variable definitions and descriptions are based off information published in the NRC's Data-Based Assessment of Research-Doctorate Programs in the US database and Methodology Guide, released in 2010–11. 'Allocated' faculty are those who supervise dissertations across multiple programs. Their position assignment (=1) is therefore split proportionally across the programs they are affiliated with so that the sum of their allocated positions was equal to 1 to any faculty member supervising dissertation. 'Core' faculty are those whose primary appointment is in the doctoral program while 'new' faculty are those with tenure track appointments who were appointed in 2003–6.

Table A4. Descriptive statistics of covariates by awardee status threshold.

	Full sample	Honorable mentions only	At least one award	
Leadership quality and composition				
Average publications per faculty	1.78	1.52	1.94	***
Range: 0.01–10.16	(1.28)	(1.11)	(1.35)	
Faculty with grants (0–100) %	78.13	76.37	79.20	**
Female faculty (0–100) %	20.27	21.36	19.61	**
Interdisciplinary faculty (0–95) %	26.16	24.98	26.88	*
Non-Asian minority faculty (0–55) %	4.38	4.47	4.32	
Peer quality and composition				
Average GRE quantitative score	726.68	714.03	734.35	***
Range: 485.83–800	(51.93)	(55.59)	(48.03)	
Female students (0–91) %	41.31	42.28	40.71	*
Non-Asian minority students (0–100) %	10.79	11.37	10.43	*
Students with academic plans (0–100) %	55.32	55.67	55.10	
Department support and traits				
Student workspace provided	0.80	0.79	0.81	
Student proposal support provided	0.69	0.69	0.69	
Number of student programs	16.50	16.37	16.57	**
Range: 0–18	(1.79)	(1.82)	(1.76)	
Median time to degree, years	5.59	5.59	5.59	
Range: 2–12	(1.05)	(1.08)	(1.03)	
Program size quartiles				
Q1, smallest	0.09	0.14	0.07	***
Q2	0.21	0.28	0.17	***
Q3	0.29	0.31	0.28	
Q4, largest	0.40	0.27	0.48	***
University traits				
Region				
Northeast	0.248	0.233	0.257	
Midwest	0.218	0.218	0.218	
South Atlantic	0.171	0.187	0.162	
South Central	0.10	0.12	0.08	***
West	0.268	0.238	0.286	**
Public university	0.65	0.75	0.60	***
University rank				
High	0.66	0.57	0.72	***
Mid	0.26	0.32	0.22	***
Low	0.08	0.11	0.07	***
Observations	1033	390	643	

Notes: Means and (standard deviations) or proportions presented; $***P < 0.01$, $**P < 0.05$, $*P < 0.1$; See Table A.3 for details on variable composition and timing; threshold for stratification: program only receives honorable mentions versus program receives at least one award.

Table A5. Alternate threshold comparison of average marginal effects for Step 1 estimations.

Variables	(1) At least one award	(2) At least two awards	(3) At least two awards
Leadership quality and composition			
Standardized average publications per faculty	**0.072*****	**0.099*****	**0.105*****
	(0.022)	(0.020)	(0.024)
Standardized percent of faculty with grants	−0.002	0.015	0.006
	(0.016)	(0.017)	(0.019)
Standardized percent female faculty	−0.011	0.022	0.021
	(0.015)	(0.016)	(0.018)
Standardized percent of interdisciplinary faculty	**0.030****	**0.026***	**0.035****
	(0.014)	(0.014)	(0.015)
Standardized percent non-Asian minority faculty	0.000	−0.003	0.005
	(0.014)	(0.015)	(0.017)
Peer quality and composition			
Standardized average GRE quantitative score	**0.067*****	**0.094*****	**0.106*****
	(0.019)	(0.021)	(0.023)
Standardized percent female students	−0.005	0.004	0.001
	(0.019)	(0.020)	(0.023)
Standardized percent of non-Asian minority students	0.013	**0.034****	**0.042****
	(0.015)	(0.016)	(0.017)
Standardized percent of students with academic plans	0.008	0.026	0.032*
	(0.015)	(0.016)	(0.017)
Program support and traits			
Student workspace provided (binary)	0.022	0.001	0.010
	(0.038)	(0.037)	(0.042)
Student proposal support (binary)	0.012	**0.087*****	**0.086****
	(0.032)	(0.031)	(0.036)
Above or below field average in student support programs (binary)	**0.083*****	0.045	**0.068****
	(0.031)	(0.031)	(0.034)
Standardized median time to degree	−0.008	0.017	0.027
	(0.016)	(0.016)	(0.019)
Program size quartiles (referent: Q1, smallest)			
Program size, Q2	0.057	**0.132****	**0.149****
	(0.052)	(0.066)	(0.068)
Program size, Q3	**0.096***	**0.147****	**0.171*****
	(0.052)	(0.064)	(0.066)
Program size, Q4 (largest)	**0.220*****	**0.287*****	**0.341*****
	(0.053)	(0.063)	(0.065)
University traits			
Public university (binary)	**−0.148*****	**−0.090****	**−0.143*****
	(0.042)	(0.039)	(0.046)
University rank (referent: low)			
University rank: high	0.021	0.021	0.046
	(0.057)	(0.062)	(0.069)
University rank: mid	−0.003	−0.042	−0.019
	(0.055)	(0.061)	(0.067)
Observations	1,028	981	705
Academic field fixed effects	Yes	Yes	Yes
Region controls included	Yes	Yes	Yes

Notes: Logistic regression results; average marginal effects presented; standard errors in parentheses; $^{***}P < 0.01$, $^{**}P < 0.05$, $^{*}P < 0.1$; all standardized covariates are normalized by academic division.

Outcomes: (1) Binary indicator in any awards received (primary model from Table 3, Column 1).

(2) Binary indicator if at least two awards received compared to zero or one.

(3) Binary indicator if at least two awards received compared to zero.

Of total sample, 390 observations with zero awards, 300 with one award, 343 with at least two awards. Due to perfect prediction, certain narrow fields were omitted from the estimations: For (1) computer science, mathematics, nutrition, and statistics and probability. For (2) and (3): applied mathematics, communication, computer sciences, entomology, forestry, mathematics, nutrition, physics, physiology, statistics and probability.

Table A6. Alternate sample comparison of average marginal effects for Step 1 estimation.

Variables	(1) Full sample	(2) No social sciences
Leadership quality and composition		
Standardized average publications per faculty	**0.072***** (0.022)	**0.074***** (0.024)
Standardized percent of faculty with grants	−0.002 (0.016)	0.004 (0.018)
Standardized percent female faculty	−0.011 (0.015)	−0.004 (0.016)
Standardized percent of interdisciplinary faculty	**0.030**** (0.014)	**0.035**** (0.016)
Standardized percent non-Asian minority faculty	0.000 (0.014)	−0.001 (0.016)
Peer quality and composition		
Standardized average GRE quantitative score	**0.067***** (0.019)	**0.049**** (0.021)
Standardized percent female students	−0.005 (0.019)	−0.003 (0.020)
Standardized percent of non-Asian minority students	0.013 (0.015)	0.013 (0.016)
Standardized percent of students with academic plans	0.008 (0.015)	0.008 (0.018)
Program support and traits		
Student workspace provided (binary)	0.022 (0.038)	0.021 (0.048)
Student proposal support (binary)	0.012 (0.032)	0.040 (0.035)
Above or below field average in student support programs (binary)	**0.083***** (0.031)	**0.091***** (0.034)
Standardized median time to degree	−0.008 (0.016)	−0.003 (0.017)
Program size quartiles (referent: Q1, smallest)		
Program size, Q2	0.057 (0.052)	0.027 (0.058)
Program size, Q3	**0.096*** (0.052)	0.082 (0.058)
Program size, Q4 (largest)	**0.220***** (0.053)	**0.203***** (0.059)
University traits		
Public university (binary)	**−0.148***** (0.042)	**−0.167***** (0.047)
University rank (referent: low)		
University rank: high	0.021 (0.057)	0.012 (0.064)
University rank: mid	−0.003 (0.055)	−0.008 (0.061)
Observations	1,028	804
Academic field fixed effects	Yes	Yes
Region controls included	Yes	Yes

Notes: Logistic regression results; average marginal effects presented; standard errors in parentheses; ***$P < 0.01$, **$P < 0.05$, *$P < 0.1$; all standardized covariates are normalized by academic division. Due to perfect prediction, certain narrow fields were omitted from the estimations: For (1) and (2) computer science, mathematics, nutrition, and statistics and probability.