Running head: UNDERGRADUATE RESEARCH PROGRAMS

Making a Difference in Science Education for Underrepresented Students: The Impact of Undergraduate Research Programs

M. Kevin Eagan Jr., Gina A. Garcia, Felisha Herrera,
Juan Garibay, Sylvia Hurtado, Mitchell J. Chang
University of California, Los Angeles

Contact: M. Kevin Eagan Jr., 405 Hilgard Ave., 3005 Moore Hall, University of California,
Los Angeles, CA 90095-1521; Phone: (310) 825-1925.

This study was made possible by the support of the National Institute of General Medical Sciences, NIH Grant Numbers 1 R01 GMO71968-01 and R01 GMO71968-05 as well as the National Science Foundation, NSF Grant Number 0757076. This independent research and the views expressed here do not indicate endorsement by the sponsors.

ABSTRACT: To increase the numbers of underrepresented minority students in science and engineering, agencies such as the National Science Foundation and National Institutes of Health have allocated significant funding to undergraduate research programs, which have been shown to increase the likelihood of continuing on to graduate school. This study utilized propensity score matching and multinomial logistic regression to examine how participation in undergraduate research affects students’ likelihood to enroll in STEM-related graduate programs. Findings indicate that participation in a structured undergraduate research program significantly improved students’ probability of indicating intentions to enroll in a STEM graduate program even after accounting for potential selection biases in the data.

Introduction

Despite the fact that graduate enrollment in science and engineering fields increased between 2000 and 2007, the number of underrepresented racial minority (URM) students entering graduate and professional programs in these disciplines continues to lag behind the number of White and Asian American students pursuing post-baccalaureate degrees in science and engineering (National Science Foundation [NSF], 2009). During the 2006-2007 academic year, Native American, Black, and Latino students represented just 0.4%, 4.9%, and 3.6%, respectively, of all science, technology, engineering, and mathematics (STEM) graduate students (Council of Graduate Schools, 2007). This ongoing troublesome phenomenon at the graduate level can be traced back to the enduring disparity in undergraduate completion rates between URM students and White and Asian American students in STEM.

A decade ago, Huang et al. (2000) found that, although URM undergraduate students and their White and Asian American counterparts initially show similar interest in pursuing STEM degrees, only 26.8% of URMs completed undergraduate degrees in STEM disciplines within five years whereas White and Asian American students had STEM completion rates of 46%. Although nearly a decade has passed since the release of the study by Huang and colleagues, a similar, and perhaps more worrisome, picture exists today. A recent report by the Higher Education Research Institute (HERI) at UCLA shows that, despite displaying similar initial interest in STEM majors, Black, Latino, and Native American students had five-year STEM bachelor’s degree completion rates of 18.4%, 22.1%, and 18.8%, respectively; in comparison, 33% of White and 42% Asian American initial STEM majors completed a bachelor’s degree in STEM within five years of entering college (HERI, 2010).

To address the gap in STEM bachelor's degree attainment and the under-representation of Black, Latino, and Native American students in STEM graduate programs, federal agencies, such as the National Institutes of Health (NIH) and the National Science Foundation (NSF), have invested significantly in undergraduate research programs geared toward retaining URM students in STEM disciplines and facilitating their matriculation into STEM graduate programs. Past research examining the effects of undergraduate research programs has concluded that these programs represent an important catalyst in encouraging students to pursue graduate study (Barlow and Villarejo, 2004) and increasing students' commitment to STEM careers and STEM graduate programs (e.g., Hunter, Laursen, & Seymour, 2007; Lopatto, 2004; MacLachlan, 2006; Russell, Hancock, & McCullough, 2007; Seymour, Hunter, Laursen, & DeAntoni, 2004), yet many of these studies have serious shortcomings.

For example, the vast majority of scholarship on the student-derived benefits from undergraduate research participation has analyzed only data collected from single institutions and individual programs, and such analytic approaches limit the generalizability of the findings to other institutions and initiatives. Furthermore, studies tend to be retrospective in nature by asking alumni from undergraduate research programs to discuss their experiences or to identify the key undergraduate opportunities that enabled them to pursue graduate school (e.g., Hurtado, Cabrera, Lin, Arellano, & Espinosa, 2009; Barlow & Villarejo, 2004; Bauer & Bennett, 2003; Hathway, Nagda, & Gregerman, 2002; MacLachlan, 2006). Other studies have used simple comparisons between undergraduate research program participants and nonparticipants in examining graduate school enrollment rates (e.g., Lopatto, 2004; Maton & Hrabowski, 2004). A third category of studies includes qualitative and quantitative longitudinal research designs to examine the influences of undergraduate research experiences over time (Russell, Hancock, &

McCullough, 2007; Hunter, Laursen, & Seymour, 2006; Seymour, Hunter, Laursen, & Deantoni, 2004). None of these studies has accounted for the selection bias inherent in the data, as students do not randomly decide to participate in undergraduate research programs.

By not accounting for the non-random assignment, or selection, of student participants in undergraduate research programs, prior studies may have misestimated the short- and long-term effects of participation in an undergraduate research program. This study aims to address this limitation by using propensity score matching to estimate the treatment effects of participation in an undergraduate research program on students' likelihood of enrolling in either a STEM graduate program or a non-STEM graduate program versus choosing not to enroll in graduate school. Propensity score matching, which we discuss more fully in the methods section, is a statistical technique that adjusts samples for potential endogeneity bias to provide a more accurate estimation of the treatment effect (Pearl, 2009). Given that federal agencies and individual institutions have invested decades of funding toward undergraduate research programs with a goal of improving the educational success of STEM students, the purpose of this study is to examine the effects of these programs on students' graduate school enrollment intentions through the use of advanced statistical techniques.

Literature Review

Graduate School Enrollment

That students extend their undergraduate studies into graduate school is widely considered to be an important outcome for sustaining our nation's science capacity. A range of factors concerning students' academic achievement, background, and college involvement has been shown to contribute to students' likelihood of enrolling in graduate or professional programs. Prior academic achievement, measured by SAT scores and undergraduate GPAs, has been identified as one of the strongest predictors of graduate school enrollment (Ekstrom, Goertz, Pollack, & Rock, 1991;Fox, 1992; Heller, 2001; Millett, 2003; Weiler, 1991). For example, Mullen, Goyette, & Soares (2003) found that scoring higher on the SATs and earning higher grades in college significantly improved students' chances of enrolling in a master's program or a doctoral program. Given that graduate school admission decisions are based upon students' academic record and their likelihood of success as a graduate student (Landrum, Jeglum, & Cashin, 1994; Purdy, Reinehr, & Swartz, 1989), the strong, positive effect between prior academic achievement and graduate school enrollment is not surprising.

Although studies have shown that students' background characteristics significantly predict graduate school aspirations (Carter, 1999, 2001), decision to apply (Perna, 2004), and enrollment (Ekstrom, Goertz, Pollack, & Rock, 1991; Perna, 2004), more research is needed to explore the factors that contribute to URM students' likelihood of graduate enrollment, particularly in STEM fields. Several studies have noted conflicting results regarding the influence of race on graduate enrollment (Heller, 2001; Millett, 2003; Nettles, 1988). For example, Millet found that Black students were significantly less likely than their White peers to enroll in their first-choice graduate or professional school, but she found no difference between White students and their Latino or Asian American peers. In contrast, Heller (2001) found that Black students were significantly more likely than their White counterparts to enroll in graduate school within four years of completing their bachelor's degree.

The effects of race on graduate enrollment may be confounded by socioeconomic status, which is often operationalized as a combination of one's parents' education and financial resources. Mullen et al. (2003), for example, found that parent education continues to influence their children's educational success beyond the bachelor's degree, and parental education affects graduate enrollment indirectly through undergraduate institution selection, educational expectations, and career values. Heller (2001) also found that students who had parents who had enrolled in post-baccalaureate educational programs had significantly higher probabilities of pursuing graduate school compared to their peers whose parents had earned a bachelor's degree or less.

Other studies, however, have shown that socioeconomic status does not necessarily affect enrollment in graduate school in direct ways. Ethington and Smart (1986), for example, applied causal modeling and found that socioeconomic status indirectly affected the decision to attend graduate school by influencing students' selection of an undergraduate institution. Ethington and Smart's findings connects to other research that shows that attending an elite private or highly selective college for undergraduate studies increases the probability of attending graduate school (Eide, Brewer, & Eherenberg, 1998; Henson, 1980; Lang, 1987) and enrolling in a doctoral programs at a major research university as well as predicts the quality of graduate schools selected (Zhang, 2005).

Beyond college quality, a higher level of academic involvement significantly and positively predicts graduate school enrollment among women and men who majored in STEM (Sax, 2001). Students who more frequently interacted with faculty and earned higher grades in college were significantly more likely to pursue a graduate degree (Sax, 2001). Additionally, student-faculty interactions have been cited as key reasons by URMs for pursuing graduate study (Carter, 2002; Ibarra, 1996; Maton, Hrabowski, & Schmitt, 2000). Given these findings, this study considers the influence of faculty support and student' connection to the peers on decisions to enroll in graduate school.

Undergraduate Research Programs

Another important source of support for STEM students and the central focus of this study, are undergraduate research programs. Although previous studies have documented the benefits of these programs, those studies generally have fallen into three categories: retrospective analyses of data collected from program alumni; cross-sectional studies that simply compare graduate enrollment rates between program participants and non-participants; and longitudinal studies using qualitative or quantitative data that, even with their longitudinal designs, ignore the inherent endogeneity of the data. For example, in a retrospective study, Bauer and Bennett (2003) surveyed alumni from a mid-sized, research-intensive institution who participated in an established Undergraduate Research Program (URP) and compared them to alumni who were not part of URP. Bauer and Bennett found that 80% of the URP alumni pursued graduate school compared to 59% of the non-research participants. Similarly, Hathaway et al. (2002) compared three groups of alumni: those who had participated in either (1) a sponsored Undergraduate Research Opportunity Program (UROP), (2) an alternative research experience on campus, and (3) no research activity during their undergraduate tenure. The authors reported that 81.5% of UROP alumni and 82% of alumni who had an alternative research experience pursued graduate education, whereas only 65.4% of alumni who did not engage in research went on to graduate school.

In another retrospective study, Barlow and Villarejo (2004) compared students who participated in the Biology Undergraduate Scholars Program (BUSP) to the general population, and found that program participants were more likely to attend graduate school. Specifically, they found that 8% of those who participated in BUSP had enrolled in a Ph.D. program within one year of completing their bachelor’s degree compared to 4% of the general population of alumni from the same campus. In a series of focus groups conducted by Hurtado et al. (2009), URM students identified undergraduate research programs as important for helping them stay committed to the sciences and for bridging undergraduate research with graduate school and professional careers in research. Maton and Hrabowski (2004) also argued that research internships and mentors are critical for leading students into Ph.D. programs.

In addition to improving graduate school enrollment, undergraduate research programs also have been linked to a stronger commitment to research careers in science. In a study aimed at understanding the academic experiences of scientists and engineers of color, MacLachlan interviewed 158 African American, Chicano, and Native American Ph.D. recipients and found that participation in sponsored research programs and other field experiences contributed to interest in research careers. A majority of the African American respondents in MacLachlan's study, for example, reported increased interest in their field through research programs and related jobs. MacLachlan concluded that for students of color, the combination of participating in a research program and interacting with a mentor improved the likelihood of pursuing graduate school.

Seymour et al. (2004) have documented why participating in an undergraduate research program might improve students' interest in pursuing graduate studies. In their longitudinal qualitative study, they found that students tend to connect their experiences in research programs with increased confidence in conducting research, defending their findings, and making contributions to their discipline. Research program participants also reported having gained a deeper level of knowledge and understanding of scientific theory and concepts as well as an increase in critical thinking and problem-solving skills through their participation.

Although a number of previous studies have revealed that STEM students who participate in undergraduate research have a higher likelihood of pursuing graduate education, those studies generally have focused on one institution or a single program. Subsequently, they have been limited by potential selection biases inherent in students' decision to participate in an undergraduate research program. Utilizing national data from the Cooperative Institutional Research Program's (CIRP) 2004 Freshman Survey and 2005 Your First College Year survey, Hurtado et al. (2008) examined individual characteristics, social aspects, and structural factors that predicted students' likelihood of participating in an undergraduate research program in their first year of college. They found that students' decisions to participate in undergraduate research programs do not occur randomly but are rather influenced by a range of factors. For example, Black students were less likely than their White peers to participate in structured research programs in the first year. Additionally, students who either intended to live on campus the first year, enrolled in a first-year experience course, or joined a pre-professional or departmental club also had higher probabilities of participating on a research project during their first year. Likewise, institutional size and the availability of a structured health science program were also related to higher participation levels.

Social Capital, Cultural Capital, and Science Identity

Those potential educational benefits associated with participating in a research program can be broadly understood through two complementary frameworks. First, those programs can be said to enhance participants' social and cultural capital (Bourdieu, 1986; Coleman, 1988, 1990; Portes, 1998) by developing their social and personal networks and sources of information, which improve students' capacity to navigate the educational system (Lin, 1999, McDonough, 1997, 1998; Swartz, 1997). According to Hurtado et al. (2008), initiatives with support systems that provide high levels of mentoring and peer relationships and acquaint students to scientific norms better enable students to access opportunities at their undergraduate institutions that will develop their science orientation because such initiatives essentially enhance students' social and cultural capital. Undergraduate research programs may well transmit such capital by providing STEM relevant direction and support to students.

Second, undergraduate research programs can also be said to develop a stronger identification with participants' respective STEM disciplines, which help orient them toward graduate and professional programs in science and engineering. According to Carlone and Johnson (2007), students' science identities can be strengthened in three key ways: (1) by fostering knowledge growth, (2) by providing opportunities to socially display scientific knowledge and practices, and (3) by building one's acknowledgement as being a "science person," especially by way of recognition by others. Research programs may well improve students' likelihood of pursuing graduate studies by developing participants' science identity in those three key ways. The present study employs the notions of both science identity and social and cultural capital as conceptual lenses to explore the impact of undergraduate programs on students' post-college expectations.

Methodology

This study addresses the following research questions:

1. What pre-college experiences and characteristics of entering college students predict their likelihood of participating in a structured undergraduate research program during college?
2. After accounting for students' chances of participating in an undergraduate research program, what effect does participation in such a program have on students' intention to enroll in graduate/professional school, particularly in a STEM field?

Sample

We address those questions by analyzing a longitudinal sample that comes from the 2004 Freshman Survey (TFS) and 2008 College Senior Survey (CSS), both of which were administered by the Cooperative Institutional Research Program (CIRP) at HERI. CIRP's TFS asked entering freshmen about their career and educational goals, high school experiences, and perspectives on a host of political and social issues. CIRP's CSS collected data about students' college experiences, future plans, and attitudes about college life and social issues.

Funding from the National Institutes of Health (NIH) and the National Science Foundation (NSF) allowed for a sampling process that targeted specific institutions and students within those institutions. We targeted colleges and universities with strong reputations for producing high numbers of STEM bachelor's degree recipients, institutions that had undergraduate research programs funded by NSF and NIH, and a number of minority-serving institutions (MSIs) that do not normally participate in CIRP-sponsored surveys. Within each institution, we identified all of the URM students who indicated plans to pursue a STEM major on the 2004 Freshman Survey. Then, we identified equivalent numbers of URM students not planning to pursue STEM and White and Asian American students who intended to major in STEM, which served as a comparison group. This sampling strategy was part of a larger study aimed at comparing URM STEM students to their same-race, different major peers and to their same-major, different-race peers. The sample for this study includes 4,212 students from all racial and ethnic backgrounds who indicated in 2004 an intention to pursue a STEM-related bachelor's degree.

Variables

The dependent variable for this study is a three-part categorical variable regarding students' intentions to pursue either a STEM-related graduate program, a non-STEM related graduate program, or a path other than graduate school. We derived this variable from a question on the CSS that asked students to report their intended graduate school major, and students who did not plan to enroll in graduate school were instructed to skip the question. Appendix A provides the list of majors that we classify as STEM, both at the graduate and undergraduate levels.

The primary independent variable of interest is whether students participated in a structured undergraduate research program during college, which is taken from an item on the 2008 CSS. Appendix B presents the full list of variables used in the analyses. To predict students' likelihood of participating in a structured undergraduate research program, we utilized a number of variables from the 2004 CIRP Freshman Survey, including race, gender, mother's education, parental income, prior academic preparation, academic self-confidence, and intended academic major. We included the following dummy variables to account for students' race: Black, Latino, Native American, and Asian American, with White serving as the reference group. Prior academic achievement was measured by students' high school GPA, composite SAT scores, and the number of years they spent studying math, physical science, and biological science in high school. We accounted for whether students had participated in a pre-college summer research program as well as students' incoming aspirations for a medical degree or a Ph.D.

Three constructs scored by CIRP using item-response theory (IRT) also were included in the model: students' academic self-concept, social self-concept, and college reputation. The items composing these factors can be found in Appendix B, and the methodology used to generate these factors is explained in more detail by Sharkness, DeAngelo, and Pryor (2010). In predicting students’ likelihood to participate in a structured undergraduate research program, we also controlled for their entering identification with the STEM major, a construct identified by Chang, Eagan, Lin, and Hurtado (in press). This construct included the following items: becoming an authority in my field, making a theoretical contribution to science, obtaining recognition from my colleagues for contributions to my special field, and working to find a cure to a health problem. Cronbach’s alpha for this factor was 0.70.

Many of the same variables used to predict students’ likelihood of participating in an undergraduate research program were also included in the analysis predicting students’ probability of enrolling in a STEM or non-STEM graduate program compared to not pursuing graduate school. Additionally, this second prediction equation included a number of college experiences. We accounted for whether students joined a club related to their major, the frequency with which they interacted with graduate students and teaching assistants, and the hours per week that they spent studying. Additionally, we included measures of students’ career goals, including desires to find a career that enabled them to have the potential for high income, achieve social recognition or status, and discover or enhance new knowledge. Students’ sense of faculty support, satisfaction with the relevance of coursework to their intended career, and cumulative GPA in college were also considered in the model. Finally, we included a control that identified whether students had persisted in a STEM-related major through four years of college.

In addition to the student-level variables, we also included a handful of institutional variables in the modeling predicting students’ graduate school decisions. We accounted for institutional control and HBCU status. Additionally, we examined the relationship between selectivity and students' likelihood to enroll in a STEM or non-STEM graduate program compared to not pursuing graduate school.

Analyses

In estimating the effect of undergraduate research program participation on students' probability to indicate intentions to enroll in a graduate STEM program or graduate program not in STEM compared to having no plans for graduate school enrollment, we relied on two analytic techniques. First, we used a logit model to estimate students' propensity, or likelihood, of participating in an undergraduate research program, and we use the propensity score generated from the model to statistically adjust the sample. This statistical adjustment enabled us to more accurately compare students who participated in a research program to peers with similar characteristics who did not participate. After statistically accounting for possible selection bias in the data, we used hierarchical multinomial logistic regression to examine how research program participation relates to students' graduate school intentions.

The issue of selection bias tends to arise in studies that rely upon analysis of ex post facto, or "after the fact," data. Titus (2007) explains that endogeneity bias "occurs when predictors of an outcome are themselves associated with other unobserved or observed variables" (p. 489). Furthermore, Titus argued that sample selection bias may result from the lack of experimental designs and not accounting for this bias may lead to inaccuracy of findings; in our case, those findings related to the effects of undergraduate research programs on post-college outcomes. Desjardins, McCall, Ahlburg, and Moye (2002) emphasized the need for higher education research to account for issues of endogeneity when examining the effects of college, or specific programs, on an array of student outcomes.

Given that our study analyzes observational data, we rely upon the counterfactual framework advanced by Rosenbaum and Rubin (1983, 1984, 1985), who expanded upon the work of scholars in the natural sciences (e.g., Fisher, 1935; Neyman, 1923, 1935). Guo and Fraser (2010) describe a counterfactual as “a potential outcome, or the state of affairs that would have happened in the absence of the cause” (p. 24). In regard to this study, for a student who participated in a research program, the counterfactual is the hypothetical outcome (graduate enrollment intention) had that student *not* taken part in an undergraduate research program. In contrast, the counterfactual for a non-participant in research is the potential probability of reporting a graduate enrollment intention if that individual *had* been a part of a research program.

Holland (1986) and Rubin (1986) suggest that, through the use of a counterfactual framework, it is possible to make causal inferences from observational data. Holland (1986) explains that no single unit (in this case, student) can be observed as both a research participant and a non-participant, “and, therefore, it is impossible to observe the effect of [a treatment] on [a unit]” (p. 947). Instead, Holland recommends comparing an individual who receives treatment (research participant) to an individual with similar characteristics who does not receive treatment (non-participant). This use of a counterfactual framework requires the estimation of a propensity score and a reweighting of the data based on estimated propensity scores.

Because treatment in our study is defined as participation in an undergraduate research program (compared to no participation), we utilized logistic regression to estimate students’ propensity to participate (Guo & Fraser, 2010). The selection model is given by:

$$P(W_i|X_i = x_i) = E(W_i) = \frac{e^{x_i \beta_i}}{1+ e^{x_i \beta_i}} \qquad (1)$$

where W_i represents the binary treatment (W_i =1 for treatment condition, W_i = for control condition) for the *i*th student (i=1,…,N), X_i represents the vector of covariates predicting selection into the treatment, and β_i represents the set of regression parameters (Guo & Fraser, 2010). In this model, an estimated propensity score of 0.45 corresponds to an individual who had a 45% probability of participating in an undergraduate research program. After generating propensity scores, we conducted analyses to determine whether our sample was balanced, and the findings from these analyses are presented in the results section.

In using propensity scores to adjust a sample, whether by matching or by reweighting, we assume that “conditional on covariates X, the assignment of study participants to binary treatment conditions (i.e., treatment vs. nontreatment) is independent of the outcome of nontreatment (Y_0) and the outcome of treatment (Y_1)” (Guo & Fraser, 2010, p. 31), which is known as the conditional independence assumption. Thus, in this study we assume that students who choose to participate in a research program expect that such participation will provide them with a benefit that is equivalent to the average benefit (i.e., probability of reporting a specific graduate school enrollment intention) for other research participants who share similar background and pre-college characteristics.

Because of limitations in the propensity score matching program (PSMATCH2) in Stata, we rely on reweighting techniques to statistically adjust the sample based on students’ likelihood of participation in an undergraduate research program (Hirano & Imbens, 2001; Nichols, 2007, 2008; Rosenbaum, 1987). In reweighting our sample, we rely on suggested calculations by Nichols (2008) and Guo and Fraser (2010) to create weights for the average treatment effect (ATE), the average treatment for the treated (ATT) effect, and the average treatment for the untreated (ATU) effect. ATE estimates the treatment effect for the entire sample whereas the ATT effect provides an estimate of the difference in an outcome between research participants and nonparticipants among individuals who had similar high probabilities of participating in a research program. In contrast, the ATU effect can be described as, among research program nonparticipants, the change in probability of reporting specific graduate school intentions *if* these individuals *had* participated in an undergraduate research program. Guo and Fraser (2010) and Nichols (2008) suggest the following calculation be used to generate a weight for the average treatment effect:

$$\omega(W,x) = \frac{W}{\hat{e}(x)} + \frac{1-W}{1-\hat{e}(x)} \tag{2}$$

where *W* corresponds to the value of treatment (1,0) and ê(x) corresponds to the propensity score. The weight for the ATT effect is given by:

$$\omega(W,x) = W + (1\text{-}W)\,\frac{\hat{e}(x)}{1-\hat{e}(x)} \tag{3}$$

where *W* corresponds to the value of treatment (1,0) and ê(x) corresponds to the propensity score (Guo & Fraser, 2010). Finally, the weight for the ATU effect is calculated by:

$$\omega(W,x) = \mathrm{W}*\frac{1-\hat{e}(x)}{\mu*\hat{e}(x)} * \frac{1}{1-\mu} + (1\text{-}\mathrm{W}) \tag{4}$$

where *W* corresponds to the value of treatment (1,0), ê(x) corresponds to the propensity score, and µ represents the proportion of students who participated in a research program (Nichols, 2008).

After creating these weights, we utilized hierarchical multinomial linear modeling, which is a special case of hierarchical generalized linear modeling (HGLM), to predict students' probability of reporting plans of pursuing a STEM graduate program or a non-STEM graduate program compared to reporting no plans for graduate school. Multinomial logistic regression is appropriate for analyses where the dependent variable has a non-ranked, categorical structure (Hosmer & Lemeshow, 2004), and using a hierarchical model accounts for the clustering effect of the data (Raudenbush & Bryk, 2002). Our outcome is a three-point categorical variable, and we have nested data, as students are clustered within institutions. By using multinomial HGLM, we partition the variance in the outcome between individuals (i.e., students) and groups (i.e., institutions) and account for the homogeneity of errors within groups (Raudenbush & Bryk). Single-level techniques, such as logistic or multinomial logistic regression, do not account for the homogeneity of errors within groups, which may lead to an underestimation of standard errors and cause analysts to erroneously conclude the significance of parameter estimates.

In building the model, we ran a fully unconditional model to examine the extent to which the average probability of reporting intentions to enroll in a graduate STEM program or a graduate non-STEM program differed across institutions. Confirming that these probabilities significantly differed, we proceeded with building the student-level model, which included all of the previously discussed student characteristics and experiences. Next, we added institutional characteristics to examine how students' institutional context, in conjunction with student experiences, related to students' probability to report intentions of enrolling in a graduate STEM or graduate non-STEM program compared to reporting no plans for graduate or professional school enrollment. We ran the final model three times using the three different weights previously described. Using Petersen's (1985) recommended calculation, we report the results from significant parameter estimates as delta-p statistics, which can be interpreted as the change in probability of reporting an intention to pursue a graduate program for every one-unit change in the independent variable.

Limitations

This study is limited in several ways. First, as with any study involving analysis of secondary data, we are limited by the survey items and their coding schemes. For example, the primary independent variable of interest, participation in an undergraduate research program, does not allow us to disentangle the quality or effects of different types of undergraduate research programs (e.g., NSF-funded, NIH-funded, institutionally based, etc.). Additionally, we rely on students' *intentions* to enroll in a specific graduate or professional program; because we lack matriculation data, we are unable to examine the relationship between research participation and actual enrollment in graduate school. Furthermore, our dependent variable does not distinguish between enrolling in a graduate program versus a professional program.

Second, although we rely on a relatively robust method to account for issues of selection bias in analyzing the effect of undergraduate research program participation on graduate school enrollment intentions, we are unable to account for a number of unobserved factors regarding why students choose to participate in these programs or how they are chosen by program administrators for participation. We control for race, prior academic achievement, and degree aspirations to account for program selection criteria, as some programs are geared toward underrepresented students or require students to maintain a certain level of academic success. Additionally, these controls aim to account for student motivation in seeking out such programs (i.e., early graduate school aspirations). The prediction equation for the propensity score does not account for students' college GPA upon entering the program; instead, we use students' high school GPA as a proxy for their academic achievement early in college.

Third, we rely on a weighting adjustment to the sample rather than using the various matching strategies described by Guo and Fraser (2010). Limitations with the program PSMATCH2 in Stata, which does not handle categorical outcomes, required the use of a reweighting method rather than one of using matching strategies like nearest-neighbor, kernel matching, or local linear regression (Guo & Fraser). Our categorical outcome necessitated the calculation of weights from propensity scores before utilizing multinomial HGLM to predict the outcome. However, we found consistent probability differences associated with research participation when we compared the weighted multinomial HGLM with the estimates from PSMATCH2 when using a dichotomous outcome (e.g., intentions for a STEM graduate program vs. no graduate school). In other words, the ATE estimate generated from the multinomial HGLM for intentions to enroll in a STEM graduate program were quite similar to the ATE estimates generated from PSMATCH2 for the dichotomous measure of intending to enroll in a graduate STEM program versus no plans for graduate school.

Findings

Table 1 includes the findings from the logistic regression predicting research participation. The results suggest that Black students were 5.71% more likely to participate in an undergraduate research program than their White peers, yet students from other racial and ethnic backgrounds were not significantly different from White students in their probability of research program participation. Students with higher levels of pre-college academic achievement, as measured by high school GPA and composite SAT scores, had significantly higher probabilities of taking part in an undergraduate research program. Specifically, a 100-point increase in composite SAT scores corresponded with a 2.27% increase in students' probability of participation. Likewise, a one-point increase on the high school GPA variable from the Freshman Survey corresponded to a 1.32% increase in an individual's likelihood of taking part in a research program.

Students who began college with prior research experience, as evidenced by participation in a pre-college summer research program, were approximately 4% more likely to join a research program in college. Additionally, students who came to college with aspirations for a Ph.D. had a 3.54% higher probability of deciding to participate in a structured undergraduate research program, but initial M.D. aspirations had no significant association with deciding to participate in an undergraduate research program. Coming to college with a stronger identification with a STEM discipline and with a greater sense of academic self-concept were also significant positive predictors of participation. Finally, the results suggest that students who intended to major in the life sciences, physical sciences, or health sciences, compared to students intending to major in nursing, were significantly more likely to participate in an undergraduate research program during college. Engineering and computer science majors were not significantly different from their peers in nursing majors in their probability of participating in an undergraduate research program.

To determine the adequacy of our statistical adjustments based on students' propensity scores, we conducted bivariate analyses that compared the characteristics of students who participated in a research program with those students who did not participate. Table 2 highlights these differences before and after the statistical adjustment based on propensity scores. Examining the findings in Table 2, we note that research participants and non-participants significantly differed ($p < 0.05$) on 19 of the 23 variables included in the logistic regression model used to predict students' likelihood of participating in a research program during college. After reweighting the sample using a derivation of students' propensity scores, we find that students no longer significantly differed on any of the 23 covariates. Importantly, we eliminated more than 90% of the bias between participants and non-participants on 19 of the 23 covariates. Other post-hoc tests for balancing indicated that we had achieved a balanced sample; therefore, we proceeded with using the previously described weights in the multinomial HGLM analyses.

Table 3 presents the delta-p statistics for the average treatment effect (ATE), the average treatment for the treated (ATT) effect, and the average treatment for the untreated (ATU) effect, and an unadjusted model regarding the relationship between undergraduate research program participation and students' probability of reporting an intention for graduate or professional school enrollment. Examining the first column of delta-p statistics indicates that participation in an undergraduate research program significantly improved students' probabilities of reporting intentions to pursue a STEM-related graduate program, as the ATE associated with this graduate school intention was 7.84%. The fact that the value of the ATT effect was slightly lower than that of the ATE suggests that students who were more likely to participate in an undergraduate research program did not enjoy a benefit that exceeded students who might be less inclined to participate in a similar program. The result pertaining to the ATU effect suggests that the benefits from undergraduate research participation on STEM graduate program enrollment intentions were slightly higher for individuals with reduced probabilities of participating in such opportunities. In other words, students who were less-inclined to participate in an undergraduate research program stand to derive a greater benefit from participation than their peers who were more likely to participate.

The findings in Table 3 indicate that not taking into account the potential endogeneity of the data may result in limited overall bias as well as self-selection bias. Overall bias can be observed as the difference between the ATE and the unadjusted model's delta-p statistics (approximately 0.5%). Self-selection bias can be determined as the difference between the unadjusted model's delta-p statistic and the ATT effect (1.47%). The findings in Table 3 show that research participation does not appear to affect students' probability of reporting intentions to pursue a *non-STEM* graduate or professional program.

Although the effect of research participation on graduate school enrollment intentions was the focus of the paper, the multinomial HGLM identified several additional variables related to students' experiences in college. Table 4 presents the results of the full model using the ATE-adjusted weight. Black and Latino students were significantly more likely than their White counterparts to report intentions to enroll in either a graduate STEM or non-STEM program, and Asian American students were significantly more likely than White students to plan to pursue a STEM-related graduate or professional program. Students who came to college with aspirations for a medical degree had a 10.73% higher probability of reporting plans for a STEM-related graduate or professional program whereas respondents who entered college with Ph.D. aspirations had an increased likelihood of reporting intentions to enroll in any graduate program regardless of discipline. The reference group for degree aspirations included bachelor's degrees, master's degrees, and J.D.s.

Spending more time interacting with graduate students and TAs during college significantly and positively predicted students' likelihood of reporting intentions to pursue a graduate program connected with STEM disciplines but had no association with plans for a non-STEM graduate program. Neither joining a club related to the academic major nor spending more time each week studying significantly predicted students' intentions to pursue STEM or non-STEM graduate programs. In contrast, students who felt supported by their faculty were significantly more likely to indicate plans for graduate school. Specifically, a one-standard deviation increase in the faculty support factor increased students' probability of reporting plans for a graduate program in a STEM discipline by 2.11% compared to an increase of 4.51% in the probability of having intentions to pursue a non-STEM graduate program.

Students' career focus also had significant associations with their graduate school plans. Students who sought a career with high social recognition and status were significantly more likely to indicate plans for a non-STEM graduate program, and this trait did not significantly predict plans for a STEM graduate program. Respondents who sought a career that allowed for the discovery and enhancement of knowledge were nearly 6% more likely to intend to enroll in a STEM graduate program; likewise, this characteristic predicted a 2.84% increase in students' probability of reporting intentions for a non-STEM post-baccalaureate degree. Earning higher grades in college significantly and positively predicted enrollment in a graduate program (either STEM or non-STEM), and, not surprisingly, persisting in a STEM-related discipline throughout their undergraduate career significantly and positively predicted students' plans for a STEM-related graduate program. In contrast, students who persisted within a STEM field throughout the four years of their undergraduate career were significantly less likely (delta-p = -26.82%) to report plans for a non-STEM graduate or professional degree.

Finally, the results in Table 4 also provide estimates for the contextual effects of the undergraduate institution that students attended. Students who were enrolled in more selective colleges and universities were significantly more likely to report plans to pursue a STEM-related graduate or professional degree. Selectivity did not significantly predict plans for a non-STEM graduate school enrollment. No other institutional characteristics had a significant association with STEM or non-STEM graduate school plans.

Discussion and Conclusions

Results from this study support previous findings regarding the benefits of undergraduate research participation (e.g., Barlow & Villarejo, 2004; Bauer & Bennett, 2003; Lopatto, 2004). Specifically, we found that, even after accounting for possible self-selection bias among participants in undergraduate research programs, initial STEM aspirants who gained research experience through these programs were significantly more likely to indicate intentions to pursue a graduate or professional degree in a STEM-related discipline compared to their peers who did not participate in these research programs. Since participation in undergraduate research programs for STEM majors does not appear to significantly affect students’ intentions to pursue graduate or professional degrees in non-STEM fields, it suggests that these programs are structured in specialized ways that specifically enhance STEM degree aspirations.

In using more robust statistical techniques to analyze a larger sample of students and institutions, we found that the effect of undergraduate research program participation on graduate school enrollment intentions is weaker than prior studies might suggest. These prior studies generally used descriptive statistics and retrospective data to compare graduate enrollment rates of research participants with those of non-participants and found enrollment differentials that ranged from 17% to 21% (e.g., Bauer & Bennett, 2003; Hathaway et al., 2002). Although we found a significant, positive effect from participation in an undergraduate research program on intentions to pursue a STEM-related graduate or professional degree, our results suggest that such participation provides a more modest benefit to students, enhancing the likelihood of pursuing a graduate program in STEM by between 7% and 8%. Importantly, however, our study examines students’ *intentions* to enroll in graduate school rather than actual enrollment; future research should consider applying similar statistical techniques on actual graduate matriculation.

Given that participation in an undergraduate research program was positively predicted by having a stronger connection to STEM, stronger academic credentials at college entry, and prior research experience, we note that these programs are attracting and admitting students who already have begun to identify themselves as scientists (Carlone & Johnson, 2007). These students may seek out these research programs because of, or perhaps despite of, their more developed STEM identity. In participating in these programs, they are further encouraged to pursue advanced studies in STEM. Put in another way, these research programs may be said to enhance the social and cultural capital of participants and subsequently further develop one’s science identity, which lead to an increased likelihood of pursuing further STEM studies. Indeed, Seymour et al.'s (2004) research claims that undergraduate research programs provide students with both additional tools to navigate the decision process related to graduate school enrollment and the necessary set of skills to make their graduate school applications more appealing.

What prior research has failed to do is to estimate the *potential* effect that participation in an undergraduate research program would have had on outcomes for *non-participants if* those non-participants had taken part in a research program. This study estimated that effect (ATU), and the findings suggest that students who *did not* participate in a research program but had similar characteristics as program participants would have derived an even stronger benefit than participants in terms of their probability to report intentions for a STEM graduate degree. This suggests that many students who stand to benefit from participating in these programs do not have access to such opportunities or are electing not to participate. Regarding the latter, we found that students with a less-developed STEM identity were less likely to pursue undergraduate research opportunities. Thus, any strategy involving expansion of programs to reach more students would need to consider both developing students' talents and connections with STEM disciplines, as well as harvesting the talents of students with strong STEM identities.

In addition to the effects of undergraduate research on students' probability to report intentions to pursue either a STEM graduate program or a non-STEM graduate program, several other college experiences had significant associations with these two categories on the outcome variable. Although prior research suggest that undergraduate research programs provide students with a mentoring experience where they can connect more meaningfully with faculty (MacLachlan, 2006), our measure of research participation did not directly assess the presence of faculty support or mentorship within the program. Our analyses, however, included a measure of students' perceived sense of faculty support, and the findings suggest that students who had a stronger sense of faculty support also tended to have higher probabilities of reporting plans for either a graduate STEM program or a non-STEM post-baccalaureate degree, which supports the results of previous studies (Carter, 2002; Ibarra, 1996; Maton, Hrabowski, & Schmitt, 2000). It appears then that feeling supported by faculty is an important type of social network that provides students with the guidance necessary for making decisions about post-baccalaureate study (Coleman, 1988, 1990). Likewise, more frequently interacting with graduate students and teaching assistants positively predicted students' likelihood of reporting plans for a STEM-related graduate or professional degree. Thus, whether it be from faculty or graduate students, higher levels of mentorship and support appear to encourage students to further develop their identity with their STEM discipline and assist them in navigating the decision-making process related to graduate school enrollment.

In conclusion, the findings of this study adds to the literature regarding the effect of structured undergraduate research programs on undergraduate and post-baccalaureate outcomes in several important ways. Our study utilized propensity score matching techniques and utilized longitudinal data collected from a national sample of more than 4,000 students attending more than 200 different colleges and universities, whereas previous studies tended to apply less-robust statistical methods and were limited to data collected from a single institution. With these strengths, we can make stronger claims about a possible causal relationship between STEM students' participation in an undergraduate research program and their likelihood of further pursuing graduate studies in STEM. The findings suggest that these structured undergraduate research programs are wise investments for governmental agencies and institutions that strive to contribute to the larger goal of sustaining our nation’s capacity to flourish in the areas of science and technology.

References

Barlow, A. E. L. & Villarejo, M (2004). Making a difference for minorities: Evaluation of an educational enrichment program. *Journal of Research in Science Teaching, 41*(9), 861-881.

Bauer, K.W., & Bennett, J.S. (2003). Alumni perceptions used to assess undergraduate research experience. *Journal of Higher Education, 74*(2), 210-230.

Bourdieu, P. (1986). The forms of Capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). New York: Greenwood.

Bourdieu, P. & Passerson, J. (1977). *Reproduction in education, society, and culture.* London: Sage.

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching, 44 (8),* 1187-1218.

Carter, D. F. (1999). The impact of institutional choice and environments on African-American and White students' degree expectations. *Research in Higher Education, 40*(1), 17-41.

Carter, D. F. (2001). *A dream deferred? Examining the degree aspirations of African American and White college students*. New York: Routledge Falmer.

Carter, D. F. (2002). College students' degree aspirations: A theoretical model and literature review with a focus on African American and Latino students. In J.C. Smart (ed) *Higher education: A handbook of theory and research.* Bronx, NY: Agathon Press.

Chang. M. J., Eagan, M. K., Lin, M. H., & Hurtado, S. (in press). Considering the Impact of Racial Stigmas and Science Identity: Persistence Among Biomedical and Behavioral Science Aspirants. *Journal of Higher Education.*

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*(S), 95-120.

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge: Belknap Press.

Council of Graduate Schools (2007). *Graduate education: The backbone of American competitiveness and innovation.* A report from the Council of Graduate Schools Advisory Committee on Graduate Education and American Competitiveness. Washington, DC: Council of Graduate Schools.

DesJardins, S. L., McCall, B. P., Ahlburg, D. A., & Moye, M. J. (2002). Adding a timing light to the “tool box.” *Research in Higher Education, 43*(1), 83-114.

Eide, E., Brewer, D. J., & Ehrenberg, R. G. (1998). Does it pay to attend an elite private college? Evidence on the effects of undergraduate college quality on graduate school attendance. *Economics of Education Review*, *17*(4) 371-376.

Ekstrom, R., Goertz, M., Pollack, J., & Rock, D. (1991). *Undergraduate debt and participation in Graduate Education: The relationship between educational debt and graduate school aspirations, application, and attendance among students with a pattern of full-time, continuous postsecondary education.* Princeton, N. J.: Graduate Record Examinations Board.

Ethington, C. A., & Smart, J. C. (1986). Persistence to Graduate Education. *Research in Higher Education, 24*(3), 287-303.

Fisher, R. A. (1935). *The Design of Experiments.* Edinburgh: Oliver & Boyd.

Fox, M. (1992). Student debt and enrollment in graduate and professional school. *Applied Economics*, *24*(7), 669-667.

Guo, S. & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications.* Los Angeles: Sage.

Hathaway, R. S., Nagda, B. A., & Gregerman, S. R. (2002). The relationship of undergraduate research participation to graduate and professional education pursuit: An empirical study. *Journal of College Student Development*, *43*(5), 614-631.

Heller, D. E. (2001). *Debts and decisions: Student loans and their relationship to graduate school and career choice.* Indianapolis, IN: USA Group Foundation. [On-line] Available: http://www.luminafoundation.org/publications/debtsdecisions.pdf

Henson, J. (1980). Institutional excellence and student achievement: A study of college quality and its impact on education and career achievement. *Dissertation Abstract International*, *41*, 958A.

Higher Education Research Institute (2010). *Degrees of success: Bachelor's degree completion rates among initial STEM majors.* Retrieved on March 1, 2010, from http://www.heri.ucla.edu/nih/HERI_ResearchBrief_OL_2010_STEM.pdf

Hirano, K., & Imbens, G. W. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology, 2*(3-4), 259-278

Holland, P. W. (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association, 81*(396), 945-970.

Hosmer, D.W., & Lemeshow, S. (2004). *Applied logistic regression.* New York: John Wiley & Sons, Inc.

Huang, G., Taddese, N., & Walter, E. (2000). *Entry and persistence of women and minorities in college science and engineering education*. Retrieved March 16, 2007, from the U.S. Department of Education, National Center for Education Statistics website: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000601

Hunter, A. B., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education, 91*(1), 36-74.

Hurtado, S., Cabrera, N. L., Lin, M. H., Arellano, L., & Espinosa, L. L. (2009). Diversifying Science: Underrepresented Student Experiences in Structured Research Programs. *Research in Higher Education, 50*(2), 189-214.

Hurtado, S., Eagan, M. K., Cabrera, N. L., Lin, M. H., Park, J., Lopez, M. (2008). Training future scientists: Predicting first-year minority student participation in health science research. *Journal of Research in Higher Education, 49*, 126-152.

Ibarra, R. A. (1996). *Enhancing the minority presence in graduate education vii: Latino experiences in graduate education: Implications for change*. Washington, DC: Council of Graduate Schools.

Kardash, C. M. (2000). Evaluation of an undergraduate research experience: Perceptions of undergraduate interns and their faculty mentors. *Journal of Educational Psychology, 92 (1),* 191-201.

Kinkead, J. (2003). Learning through inquiry: An overview of undergraduate research. *New Directions for Teaching and Learning, 93*, 5-17.

Landrum, R. E., Jeglum, E. B., & Cashin, J. R. (1994). The decision-making processes of graduate admissions committees in psychology. *Journal of Social Behavior and Personality, 9*, 239-248.

Lang, D. (1987). Stratification and prestige hierarchies in graduate and professional education. *Sociological Inquiry, 57*(1), 12-31.

Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): First findings. *Cell Biology Education, 3*(4), 270-277.

Lin, N. (1999). Social networks and status attainment. *Annual Review of Sociology, 25*(1), 467-487.

MacLachlan, A. J. (2006). Developing graduate students of color for the professoriate in science, technology, engineering, and mathematics (STEM). *Research and Occasional Paper Series: CSHE.6.06*, Center for Studies in Higher Education, University of California, Berkeley.

Maton, K. I., Hrabowski, F. A., & Schmitt, C. L. (2000). African American college students excelling in the sciences: College and postcollege outcomes in the Meyerhoff Scholars Program. *Journal of Research in Science Teaching*, *37*(7) 629-654.

McDonough, P. M. (1997). *Choosing Colleges: How social class and schools structure opportunity*. Albany, NY: SUNY Press.

McDonough, P. M. (1998). Structuring college opportunities: A cross-case analysis of organizational cultures, climates, and habiti. In C. A. Torres & T. R. Mitchell (Eds.), *Sociology of education: Emerging perspectives* (pp. 181-210). Albany: SUNY Press.

Millett, C. M. (2003). How undergraduate loan debt affects application and enrollment in graduate or first professional school. *The Journal of Higher Education, 74*(4), 386-427.

Mullen, Ann L., Goyette, K. A., & Soares, J. A. (2003). Who goes to graduate school? Social and academic correlates of educational continuation after college. *Sociology of Education*, *76*(2), 143-169.

National Science Foundation, Division of Science Resources Statistics. (2009). *S&E graduate enrollments accelerate in 2007; enrollments of foreign students reach new high* (NSF 09-314). Washington, DC: National Science Foundation.

Nettles, M. T. (1988). *Toward Black undergraduate student equality in American higher education.* Westport, CT: Greenwood Press.

Neyman, J. (1923 [1990]). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles, Section 9. D. M. Sabrowska and T.P. Speed (trans. and eds.) from the Polish original, which appeared in Roczniki Nauk Rolniczych Tom X (1923), 1-51 (Annals of Agriculture). *Statistical Science, 5,* 465-472.

Neyman, J. (1935), with co-operation by K. Iwaszkiewicz and S. Kolodziejczyk. Statistical Problems in Agricultural Experimentation (with discussion). *Supplement to the Journal of the Royal Statistical Society, 2,* 107-180.

Nichols, A. (2007). Causal inference with observational data. *The Stata Journal, 7*(4), 507-541.

Nichols, A. (2008). Erratum and discussion of propensity-score reweighting. *The Stata Journal, 8*(4), 532-539.

Pascarella, E. T., Pierson, C. T., Wolniak, G. C., and Terenzini, P. T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *Journal of Higher Education*, *75*(3), 249–284.

Pearl, J. (2009). Remarks on the method of propensity score. *Statistics in Medicine, 28*(9), 1415-1416.

Perna, L. W. (2004). Understanding the decision to enroll in graduate school: Sex and racial/ethnic group differences. *Journal of Higher Education, 75*(5), 487-527.

Petersen, T. (1985). A comment on presenting results from logit and probit models. *American Sociological Review, 50,* 130-131.

Portes, A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology, 24*(1), 1-24.

Price, D. V. (2004). *Borrowing inequality: Race, class, and student loans.* Boulder, CO: Lynne Rienner Publishers, Inc.

Purdy, J. E., Reinehr, R. C., & Swartz, J. D. (1989). Graduate admissions criteria of leading psychology departments. *American Psychologist, 44,* 960-964.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Thousand Oaks, CA: Sage Publications, Inc.

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association, 82*, 387-394.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 51-55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing Bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*(387), 516-524.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a comparison group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39,* 33-38.

Rubin, D. B. (1986). Which ifs have causal answers? *Journal of the American Statistical Association, 81*(396), 961-962.

Russell, S. H., Hancock, M.P., McCullough, J. (2007). The pipeline: Benefits of undergraduate research experiences. *Science*, 316(5824), 548-549.

Sax, L. J. (2001). Undergraduate science majors: Gender differences in who goes to graduate school. *The Review of Higher Education, 24*(2), 153-172.

Seymour, E., Hunter, A. B., Laursen, S. L., & DeAntoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three-year study. *Science Education, 88(4)*, 493-534.

Sharkness, J., DeAngelo, L., & Pryor, J. (2010). *CIRP Construct Technical Report*. Retrieved March 1, 2010 from http://www.gseis.ucla.edu/heri/PDFs/technicalreport.pdf

Stanton-Salazar, R. D. & Dornbusch, S. M. (1995). Social Capital and the Reproduction of Inequality: Information Networks among Mexican-origin High School Students. *Sociology of Education, 68*(2), 116-135.

Swartz, D. (1997). *Culture and power: The sociology of Pierre Bourdieu.* Chicago: University of Chicago Press.

Titus, M. A. (2007). Detecting Selection Bias, Using Propensity Score Matching, and estimating treatment effects: An applications to the Private Returns to a Master's Degree. *Research in Higher Education, 48*(2), 487-521.

Weiler, W. C. (1991). The effect of undergraduate student loans on the decision to pursue postbaccalaureate study. *Educational Evaluation and Policy Analysis*, *13*(3), 212-220.

Yosso, T. J. (2006). *Critical race: Counterstories along the Chicana/Chicano educational pipeline.* New York: Routledge.

Zhang, L. (2005). Advance to Graduate Education: The effect of college quality and undergraduate majors. *The Review of Higher Education, 28*(3), 313-338.

Appendix A

Science, Technology, Engineering, & Math (STEM) Majors

1. General Biology
2. Biochemistry/Biophysics
3. Botany
4. Environmental Science
5. Marine (Life) Science
6. Microbiology/Bacterial Biology
7. Zoology
8. Other Biological Science
9. Aeronautical/Astronautical Engineering
10. Civil Engineering
11. Chemical Engineering
12. Computer Engineering
13. Electrical Engineering
14. Industrial Engineering
15. Mechanical Engineering
16. Other Engineering
17. Astronomy
18. Atmospheric Science
19. Chemistry
20. Earth Science
21. Marine Science
22. Mathematics
23. Physics
24. Statistics
25. Other Physical Science
26. Health Technology
27. Medicine/Dentistry/Veterinary Medicine
28. Nursing
29. Pharmacy
30. Agriculture
31. Computer Science

Appendix B
Table of Measures

Variables	Coding
Student Characteristics	
Race: Native American	1=yes, 0=no
Race: Latino	1=yes, 0=no
Race: Black	1=yes, 0=no
Race: Asian American	1=yes, 0=no
Sex: Female	1=yes, 0=no
Composite SAT score	Continuous
High School GPA	1=D through 8=A or A+
HS years of studying math	1=None to 7=Five or more
HS years of studying physical science	1=None to 7=Five or more
HS years of studying life science	1=None to 7=Five or more
Income	1=Less than $10,000 to 14=$250,000 or more
Mother's education	
Participated in a pre-college summer research program	1=yes, 0=no
2004 degree aspiration: Ph.D.	1=yes, 0=no
2004 degree aspiration: M.D.	1=yes, 0=no
Identification with STEM discipline	Factor consisting of four items: becoming an authority in my field, making a theoretical contribution to science, obtaining recognition from my colleagues for contributions to my special field, and working to find a cure to a health problem
Academic self-concept	Construct consisting of students' self-ratings for: academic ability, drive to achieve, mathematical ability, and intellecutal self-confidence
Social self-concept	Construct consisting of students' self-ratings for: leadership ability, public speaking ability, social self-confidence, and popularity
College reputation construct	Construct consisting of why students chose a particular institution: academic reputation, graduates gain admission to top graduate/professional schools, and graduates get good jobs
2004 major: Life sciences	1=yes, 0=no

2004 major: Physical sciences	1=yes, 0=no
2004 major: Health sciences	1=yes, 0=no
2004 major: Engineering or computer science	1=yes, 0=no
Joined a club related to major	1=yes, 0=no
Interacted with graduate students/TAs	1=Not at all to 3=Frequently
Career Focus: High income potential	1=Not important to 4=Essential
Career Focus: Social recognition/status	1=Not important to 4=Essential
Career Focus: Discovery/enhancement of knowledge	1=Not important to 4=Essential
Hours per week spent studying/doing homework	Ordinal: 1=none through 8=20 or more
Satisfaction with courses in major field	1=Very dissatisfied to 5=very satisfied
Satisfaction with relevance of coursework to future plans	1=Very dissatisfied to 5=very satisfied
Sense of faculty support	Continuous
College GPA	Ordinal: 1=D through 8=A or A+
Stayed in STEM major through 2008	1=yes, 0=no
Institutional Characteristics	
Control: Private	Dichotomous: 1=yes, 0=no
HBCU	Dichotomous: 1=yes, 0=no
Selectivity	Continuous
Proportion of non-White students	Continuous

Table 1

Logistic Regression Predicting Participation in an Undergraduate Research Program

	Log Odds	S.E.	Sig.	Delta-P
Native American	0.16	0.10		
Latino	0.06	0.07		
Black	0.33	0.07	***	5.71%
Asian American	0.13	0.07		
Sex: Female	-0.07	0.05		
Composite SAT score	0.00	0.00	***	2.27%
High school GPA	0.08	0.03	***	1.32%
HS years of studying math	0.04	0.05		
HS years of studying physical science	-0.02	0.02		
HS years of studying life science	-0.01	0.02		
Income	-0.02	0.01		
Mother's education	0.02	0.01		
Participated in a pre-college summer research program	0.24	0.07	***	4.03%
2004 degree aspiration: Ph.D.	0.21	0.06	***	3.54%
2004 degree aspiration: M.D.	0.09	0.07		
Identification with STEM discipline	0.05	0.01	***	0.87%
Academic self-concept	0.01	0.00	*	0.12%
Social self-concept	0.00	0.00		
College reputation construct	0.00	0.00		
2004 major: Life sciences	0.41	0.11	***	7.34%
2004 major: Physical sciences	0.58	0.12	***	10.77%
2004 major: Health sciences	0.28	0.11	**	4.90%
2004 major: Engineering or computer science	0.22	0.11		
Constant	-4.95	0.40		

Source: Logistic regression of 2004 CIRP Freshman Survey data and 2008 College Senior Survey data. Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2

Comparison of Conditional Variables before and after Adjusting the Sample with Propensity Score Weights

	Unweighted					Weighted					
	Research Participant	Non-Participant	Pct. Bias	T	p	Research Participant	Non-Participant	Pct. Bias	Pct. Bias Reducation	T	p
Sex: Female	1.58	1.64	-11.30	-2.95	0.00	1.58	1.60	-2.50	77.70	-0.51	0.61
Native American	0.06	0.05	3.70	0.99	0.32	0.06	0.06	0.00	99.20	-0.01	1.00
Asian American	1.16	1.13	8.20	2.17	0.03	1.16	1.16	1.10	86.50	0.22	0.83
Black	0.19	0.17	4.10	1.08	0.28	0.19	0.19	-0.40	89.90	-0.08	0.93
Latino	0.19	0.23	-9.40	-2.39	0.02	0.19	0.19	-0.30	96.40	-0.07	0.94
HS years studying math	6.08	5.99	16.50	4.24	0.00	6.08	6.08	0.40	97.70	0.08	0.94
HS years studying physical science	4.15	3.99	12.10	3.19	0.00	4.15	4.11	2.80	77.00	0.57	0.57
HS years studying biological science	3.85	3.78	6.30	1.64	0.10	3.85	3.84	0.40	94.00	0.08	0.94
Composite SAT (100)	12.78	11.84	53.90	13.87	0.00	12.78	12.74	1.70	96.80	0.36	0.72
Summer research participation	1.19	1.11	24.00	6.72	0.00	1.19	1.19	0.00	100.00	0.00	1.00
High school GPA	7.32	6.87	39.30	9.75	0.00	7.31	7.28	3.50	91.20	0.80	0.43
Income	8.69	8.44	7.80	2.06	0.04	8.69	8.70	-0.30	95.70	-0.07	0.95
Mother's education	5.71	5.29	21.70	5.55	0.00	5.71	5.74	-1.50	93.00	-0.32	0.75
Life sciences major	0.38	0.30	16.30	4.28	0.00	0.38	0.38	0.70	96.00	0.13	0.90
Aspiration for a medical degree	0.33	0.28	10.30	2.71	0.01	0.33	0.33	-0.20	98.00	-0.04	0.97
Academic self-concept	55.12	51.78	42.30	11.05	0.00	55.11	54.94	2.20	94.90	0.45	0.65
Social self-concept	48.50	47.26	12.90	3.37	0.00	48.49	48.53	-0.40	97.30	-0.07	0.94
College reputation factor	50.53	49.66	11.70	3.03	0.00	50.52	50.49	0.40	96.20	0.09	0.93
STEM identification	10.78	9.88	33.40	8.77	0.00	10.77	10.73	1.60	95.20	0.33	0.74
Engineering/Computer science major	0.24	0.27	-6.80	-1.75	0.08	0.24	0.24	0.40	93.80	0.09	0.93
Health sciences major	0.20	0.23	-7.60	-1.95	0.05	0.20	0.20	-0.20	97.40	-0.04	0.97
Aspiration for a Ph.D.	0.35	0.24	26.30	7.09	0.00	0.35	0.35	0.30	98.80	0.06	0.95
Physical sciences major	0.15	0.08	21.90	6.18	0.00	0.15	0.15	0.80	96.30	0.15	0.88

Source: Mean comparison test of 2004 CIRP Freshman Survey and 2008 College Senior Survey data before and after statistically adjusting the sample with weights calculated from estimated propensity scores.

Table 3

Estimated Treatment Effects for Research Participation on Each Category of the Graduate School Enrollment Intentions

	Intend to Enroll in a STEM Graduate/Professional Program				Intend to Enroll in a non-STEM Graduate/Professional Program			
	Delta-P	Log odds	S.E.	Sig.	Delta-P	Log Odds	S.E.	Sig.
Average treatment effect (ATE)	7.84%	0.39	0.16	*	4.96%	0.23	0.17	
Average treatment for the untreated (ATU)	7.95%	0.40	0.16	*	5.98%	0.28	0.18	
Average treatment for the treated (ATT)	6.91%	0.34	0.15	*	-0.45%	-0.02	0.15	
Unadjusted model	8.38%	0.42	0.14	**	1.77%	0.08	0.15	

Source: Multinomial HGLM analyses of adjusted and un-adjusted data from the 2004 CIRP Freshman Survey and 2008 College Senior Survey. Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4

Full Results from the Multinomial HGLM Estimating Enrollment in a Graduate STEM Program or a Graduate non-STEM Program

	STEM Graduate Degree				Non-STEM Graduate Degree			
	Log Odds	S.E.	Sig.	Delta-P	Log Odds	S.E.	Sig.	Delta-P
Institutional Characteristics								
Intercept	-3.85	0.67	***		0.60	0.63		
Control: Private	-0.15	0.16			-0.11	0.15		
HBCU	0.18	0.44			0.59	0.46		
Selectivity	0.13	0.06	*	2.04%	0.07	0.07		
Proportion of non-White students	-0.03	0.03			-0.02	0.03		
Student Characteristics								
Native American	0.19	0.24			0.07	0.21		
Asian American	0.39	0.17	*	5.86%	0.05	0.17		
Black	0.60	0.18	***	8.35%	0.40	0.17	*	7.69%
Latino	0.29	0.14	*	4.42%	0.33	0.15	*	6.44%
Sex: Female	-0.22	0.13			0.18	0.13		
SAT composite	0.00	0.00			0.00	0.00		
High school GPA	-0.01	0.05			-0.02	0.05		
Parental income	-0.04	0.02			-0.01	0.02		
Mother's education	0.08	0.03	*	1.32%	0.05	0.03		
2004 degree aspiration: Medical degree	0.82	0.15	***	10.73%	0.25	0.15		
2004 degree aspiration: Ph.D./Ed.D.	0.63	0.14	***	8.71%	0.48	0.13	***	8.93%
2004 major: Physical science	-0.05	0.23			0.81	0.25	***	13.93%
2004 major: Life science	-0.02	0.19			0.39	0.21		
2004 major: Health science/pre-med	0.39	0.21			0.51	0.20	*	9.46%
2004 major: Engineering/Computer science	-0.74	0.22	***	-14.87%	0.47	0.21	*	8.89%
2004 identification with STEM	0.04	0.03			-0.01	0.02		
College reputation factor	0.01	0.01			0.01	0.01		
2004 Academic self-concept	-0.01	0.01			-0.01	0.01		
Participated in a structured research program	0.39	0.16	*	5.88%	0.23	0.17		
Joined a club related to major	0.14	0.10			0.00	0.10		
Interacted with graduate students/Tas	0.25	0.09	**	3.88%	0.08	0.08		
Career focus: High income potential	-0.09	0.08			-0.09	0.08		
Career focus: Social recognition/status	0.05	0.07			0.16	0.07	*	3.15%
Career focus: Discovery/enhancement of knowledge	0.38	0.07	***	5.72%	0.14	0.06	*	2.84%
Hours per week: Studying/homework	0.07	0.04			-0.01	0.04		
Satisfaction with courses in major field	-0.06	0.08			-0.15	0.07	*	-3.14%
Satisfaction with relevance of coursework to future plans	0.10	0.06			0.06	0.06		
Sense of faculty support	0.13	0.06	*	2.11%	0.23	0.06	***	4.51%
Cumulative college GPA	0.16	0.04	***	2.60%	0.08	0.04	*	1.71%
Stayed in STEM major through 2008	1.93	0.13	***	17.59%	-1.12	0.11	***	-26.82%

Source: Multinomial HGLM analyses of adjusted (ATE weight) data from the 2004 CIRP Freshman Survey and 2008 College Senior Survey. Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$