Research in Higher Education, Vol. 45, No. 4, June 2004 (© 2004)

ETHNIC AND GENDER DIFFERENCES IN SCIENCE GRADUATION AT SELECTIVE COLLEGES WITH IMPLICATIONS FOR ADMISSION POLICY AND COLLEGE CHOICE

Frederick L. Smyth*,** and John J. McArdle*

Using Bowen and Bok's data from 23 selective colleges, we fit multilevel logit models to test two hypotheses with implications for affirmative action and group differences in attainment of science, math, or engineering (SME) degrees. Hypothesis 1, that differences in precollege academic preparation will explain later SME graduation disparities, was fully supported with respect to the outcome gap between Whites and underrepresented minorities, partially supported for that between Asians and underrepresented minorities, and between men and women. Hypothesis 2, that college selectivity, after accounting for student characteristics, will be positively associated with SME persistence, was not supported. We demonstrate that the significance of the selectivity effect is overestimated when unilevel models are used. Admission officials are advised to carefully consider the relative academic preparedness of science-interested students, and such students choosing among colleges are advised to compare their academic qualifications to those of successful science students at each institution.

KEY WORDS: affirmative action; college selectivity; engineering; gender; mathematics; minorities; SAT; science.

INTRODUCTION

> If we are to truly alleviate the problems of an inadequately educated populace and a projected shortage of scientists and engineers, we must demand that no college student be allowed to leave science *without a struggle.* (from J. P. Schaefer's introduction to Tobias, 1990)

Demand is high for citizens with college-level training in scientific, engineering, and mathematical fields (SME). Their skills are seen as vital to American

*University of Virginia, Department of Psychology.

**Address correspondence to: Frederick L. Smyth, Department of Psychology, University of Virginia, 102 Gilmer Hall, PO Box 400400, Charlottesville, VA 22904-4400. E-mail: fsmyth@virginia.edu

0361-0365/04/0600-00353/0 © 2004 Human Sciences Press, Inc.

health, economic and security interests, and careers requiring these skills are disproportionately judged among the most prestigious (Leslie, McClure, and Oaxaca, 1998; Nakao and Treas, 1990; National Academy of Sciences [NAS], 1987; National Science Board [NSB], 2002; Schoenberger, 1988; Tobias, 1990). Underrepresentation in these fields of women, American Indians, Blacks, and Hispanics, therefore, has long been cause for concern, indicative of lost personal and group opportunity, inhibition of national productivity and of the advancement of science, generally (Chipman and Thomas, 1987; Leslie et al., 1998; NAS, 1987; Oakes, 1990). Though college-level attrition from science is the tip of an attrition iceberg (far greater losses occur in the years before traditional college age, see Chipman and Thomas, 1987; Meece, Parsons, Kaczala, Goff, and Futterman, 1982; Leslieet al., 1998; Oakes, 1990), retaining college students has figured prominently in policy objectives because such students tend to be highly able (Astin and Astin, 1993; Green, 1989; Leslie et al., 1998; NAS, 1987) and their interest has weathered the many precollege pressures to pursue a less demanding course. Still, more than half of these students will not persist in SME, with greater losses among underrepresented minorities and women (Astin and Astin, 1993; Culotta and Gibbons, 1992; Green, 1989; Leslie and Oaxaca, 1998; NAS, 1987; NSB, 2002; National Science Foundation [NSF], 1992; Oakes, 1990; Seymour and Hewitt, 1997). This study is focused on ethnic and gender differences in science persistence at selective colleges and competing ideas about how such differences may be related to affirmative action admission policies.

LITERATURE REVIEW

Institutional Selectivity and Affirmative Action

Among seven key reasons for the failure of many programs designed to increase minority representation in SME, Culotta and Gibbons (1992) cite the recruitment of inadequately prepared underrepresented minority students. They quote the biology department chairman at North Carolina Central University, a historically black institution, who faults the recruitment approach of majority White institutions:

> The way we see it, the majority schools are wasting large numbers of good students. They have black students with admissions statistics [that are] very high, tops. But these students wind up majoring in sociology or recreation or get wiped out altogether. (p. 1218)

Elliott, Strenta, Adair, Matier, and Scott (1995) concluded that affirmative action admission policies at selective colleges inadvertently derail disproportionate numbers of talented minority students from the science track. They found that large ethnic differences in SME graduation rates among students initially intending SME at four elite institutions (e.g., 34% among Blacks, 61% among Whites) were accounted for by large differences on measures of high school academic preparation (e.g., greater than 1.5 *SD*s on standardized tests). The latter differences, they argue, result from affirmative action admissions. Elliott et al. suggest that relatively less well-prepared students at these institutions—regardless of ethnicity—would have had a better chance of completing SME majors at less competitive colleges. Figure 1 is a graph of data from 11 other private colleges that Elliott et al. cite as support for this inference. Though mean SAT math score (SATM) varies considerably across the colleges, the proportions of science degrees awarded to students in the top, middle, and bottom thirds of each institution's SATM distribution are similar, "about 54%, 31%, and 15%" (p. 35), respectively. Elliott et al. conclude

> a student with an SATM score of 580 who wants to be in science will be three or four times more likely to persist at institutions J and K, where he or she is competitive, than at institutions A and B, where he or she is not. (p. 35)

This relationship holds, they contend, despite the higher overall proportion of students earning science degrees at the more selective institutions. Their inference parallels the "frog pond" social comparison theory that an individual's relative standing in a group with respect to some attribute may exert greater influence on self-assessments than some absolute measure of standing on the

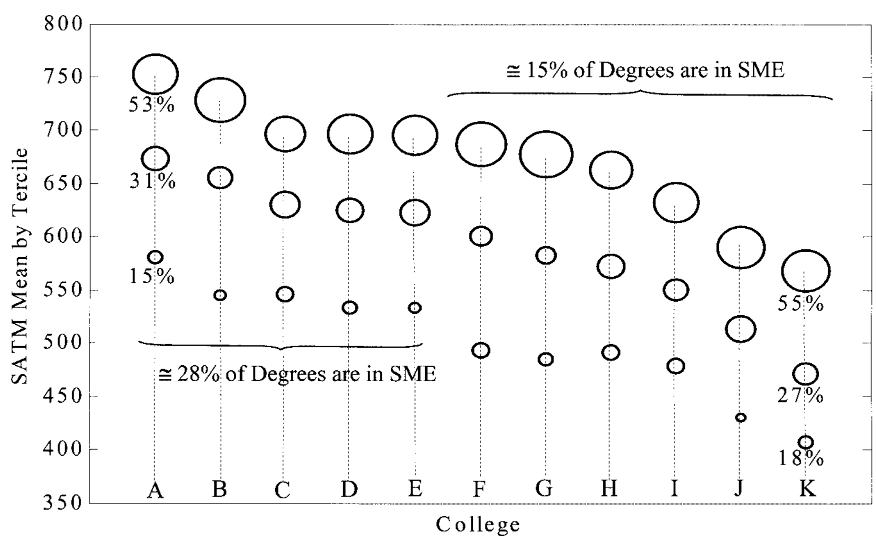

FIG. 1. For each of 11 private colleges, 3 bubbles index percentages of the degrees awarded in the natural sciences earned by students in each third of the within-institution SATM distribution (data from Elliott et al., 1995).

attribute (Buunk and Ybema, 1997; Davis, 1966; Kelley, 1952; Marsh, Kong, and Hau, 2000; Pascarella and Terenzini, 1991). Davis invoked this theory with respect to within-college effects on the career choices of male college graduates:

> Counselors and parents might well consider the drawbacks as well as the advantages of sending a boy to a 'fine' college, if, when doing so, it is fairly certain he will end up in the bottom ranks of his graduating class. (p. 31)

Bowen and Bok (1998) reached a different conclusion about affirmative action, based on higher overall graduation rates for minority students at selective colleges (28 moderately to extremely selective institutions). Despite also finding a strong negative association of selectivity with a student's collegiate rank-in-class and a greater ethnic gap in class rank at more selective schools (consistent with the frog pond theory), their positive conclusion about the results of race-sensitive admission hinges on graduation rates:

> The fact that graduation rates increase as the selectivity of the college rises and that students of the same academic ability graduate at higher rates when they attend more selective institutions shows that carefully chosen minority students have not suffered from attending colleges heavily populated by White and Asian American classmates with higher standardized test scores. Quite the contrary—they have fared best in such settings. (p. 88)

With respect to academic major, Bowen and Bok's report differs from that of Elliott et al. (1995): "Blacks and Whites were equally likely to have majored in philosophy, economics, the natural sciences, and engineering" (Bowen and Bok, p. 71). Unlike Elliott et al., however, initial intended major was not considered, and the within-ethnicity distribution of final majors was calculated only among graduates, not as a function of all matriculants. Bowen and Bok do not report on the distributions of majors as a function of institutional selectivity.

Reviews of the literature on college-level effects suggest that little value is added beyond what is explained by the preexisting attributes of the students. Toutkoushian and Smart (2001) assert, "the majority of studies . . . conclude that institutional effects contribute little, if anything, to student growth after controlling for student background and acquired characteristics" (p. 40). Pascarella and Terenzini (1991) conclude that the effect of institutional selectivity, "net" of student characteristics, "tends to be positive," but small, "no more than one or two percent of the total variance in educational aspirations, persistence, bachelor's degree attainment, and educational attainment generally" (p. 376).

Ethnic and Gender Differences

Ethnic and gender differences in college SME attainment have been often studied and are well documented (Astin and Astin, 1993; Culotta and Gibbons, 1992; Green, 1989; Leslie et al., 1998; NAS, 1987; NSB, 2002; NSF, 1992, 1999; Oakes, 1990; Seymour and Hewitt, 1997). Based on data from the *Cooperative Institutional Research Program* for $N > 26{,}000$ students entering 4-year colleges in 1985, Astin and Astin report disproportionate losses from SME majors for minority students. Persistence rates after 4 years were 37% for Chicanos, 47% for African Americans, 51% for American Indians, 61% for Whites, and 68% for Asians. More recent data (NSF, 1999) suggest that early 1990s growth in the within-group proportions of Blacks and Hispanics completing engineering and natural sciences degrees was less than that of Asians and Whites. Shortfalls among Blacks are evident despite a consistent finding that those beginning college are as, or more, likely than Whites to initially intend a SME major (Dunteman, Wisenbaker, and Taylor, 1979; Green, 1989; Lee, 1987; Leslie et al., 1998; NAS, 1987; Oakes, 1990; Post, Stewart, and Smith, 1991). Though American Indians are less frequently studied, Culotta and Gibbons (1992) conclude that they are "just as underrepresented as Blacks and Hispanics" in the scientific workforce. Asian students, on the other hand, are typically found to be both more likely to express interest in SME and to persist in the field (Culotta and Gibbons, 1992; Fullilove and Treisman, 1990; Simpson, 2001).

Though the male/female disparity varies by SME field, overall, women are less well represented than men, a finding that holds in both highly selective and more nationally representative samples (NSF, 1999; Ware, Steckler, and Leserman, 1985; Strenta, Elliott, Adair, Matier, and Scott, 1994). In the early 1990s, White men earned about 55% of the bachelor's degrees awarded in engineering and natural sciences, compared with the roughly 25% earned by White women (NSF, 1999). This disparity held within other major ethnic groups, except among Blacks, for whom the ratio was about 1:1. Furthermore, between 1990 and 1994, the gender gap widened slightly among Asians, Hispanics, and Whites, though it diminished somewhat among American Indians and Blacks.

Precollege Academic Preparation

High school grades and math test scores weigh heavily in equations predicting SME persistence, regardless of ethnicity, gender, or type of college, including single-sex female institutions (Civian and Schley, 1996; Jackson, Gardner, and Sullivan, 1993; Jagacinski, LeBold, and Salvendy, 1988; Oakes, 1990). Moreover, the Elliott et al. (1995) finding that differences in preparation account for Asian/White vs. underrepresented minority differences in science persistence is not new. Adair (1991), Astin and Astin (1993), Dunteman et al., (1979), Hilton, Hsia, Solorzano, and Benton (1989), Simpson (2001), and Ware and Lee (1988) all found that ethnic disparities in college SME persistence were not statistically significant when standard measures of academic preparation were taken into account. Astin and Astin (1993) assert, "The strongest and most consistent predictor of changes in students' interest in science majors or careers is the students' entering level of mathematical and academic competency" (p. 2). Fullilove and Treisman (1990) contend that the roots of the extraordinary Asian persistence lie in early mathematics preparation. Conversely, underrepresented minority students are consistently found to have significantly lower means on such precollege academic measures (Bowen and Bok, 1998; Dunteman et al., 1979; Elliott et al., 1995; Fullilove and Treisman, 1990; Ramist, Lewis, and McCamley-Jenkins, 1994). Gender differences favoring males are often similarly, though generally not as fully, accounted for by differences in precollege and college academic measures (Adair, 1991; Astin and Astin, 1993; Lee, 1987; Strenta et al., 1994; Turner and Bowen, 1999). It is not unusual for researchers to identify a remaining direct effect of gender, usually when predicting persistence in certain sub-categories of SME (Dunteman et al., 1979; Lee, 1987; Levin and Wyckoff, 1995; Simpson, 2001; Strenta et al., 1994).

OVERVIEW OF THIS STUDY

We use the data studied by Bowen and Bok (1998) to test the competing hypotheses about relations between SME persistence, affirmative action admission, and college selectivity. Hypothesis 1, suggested by the conclusions of Elliott et al. (1995), is that among students initially intending a SME major, within-institution differences in precollege academic preparation will account for ethnic and gender differences in SME graduation rates. Hypothesis 2, an extension of Bowen and Bok's (1998) inference about effects on overall graduation, is that college selectivity, after accounting for student characteristics, will be positively associated with SME persistence.

Our methodological goal is to advance the extant SME-persistence literature by using multilevel or hierarchical linear models (HLM) to better account for the nonrandom distribution or "nesting" of students within colleges (Burstein, 1980; Ethington, 1997; Raudenbush and Bryk, 2002; Snijders and Bosker, 1999). All of the multicollege SME studies we found in a search of the psychological and educational literature (Astin and Astin, 1993; Dunteman et al., 1979; Elliott et al., 1995; Hilton et al., 1989; Strenta et al., 1994; Ware and Lee, 1988) employed a unilevel (student-level-only) approach. That is, even when college-level variables were obtained—for example, selectivity—they were treated as characteristics of students, by assigning to each student his or her college's value on the variable. Ethington (1997) notes that such disaggregation of "higher order variables to the individual level violates the assumption of independence of observations that is a basic assumption for the classical OLS approach . . . and results in misestimated standard errors" (p. 167). The HLM approach, in which the effects of student-level variables can be estimated as a function of the corresponding mean and variation at the college level, allows for the appropriately weighted partitioning of outcome variance into its college- and student-level components.

METHOD

Participants

The participants in this study are self-identified American Indian, Asian, Black, Hispanic or non-Hispanic White students[1] who matriculated as college freshmen in 1989 at 23 colleges in the *College and Beyond* (C&B) database assembled by The Andrew W. Mellon Foundation. In addition to information from students' transcripts, the data includes self-report answers (for those at 23[2] of the 34 C&B institutions) to the national *Cooperative Institutional Research Program* (CIRP) survey (Higher Education Research Institute, 1989) administered during freshman orientation. Our analysis is limited to students at the 23 CIRP-participating colleges because the survey contained two items central to our investigation: intended academic major and an indicator of high school grades. These institutions are quite selective by national standards, more than half rated among the "most" selective in the country by Barron's Educational Guides (2001). Sixty-six percent of matriculants had complete data on key variables: high school grade average (HSGA), Scholastic Aptitude Test (SAT) scores, and records of initial intended major and graduation status. We think results of incomplete data analyses[3] support an assumption that this sample is representative of all matriculants at these colleges. Our focus will be on the 29% ($N = 5{,}047$[4]) that initially intended a SME major.[5] Their characteristics are listed by gender and ethnicity in Table 1.

Though persistence among this SME-declared group of students is our substantive concern, there were gender and ethnic differences in the likelihood of self-reporting a SME intent that we think warrant notice. Figure 2 is a plot of the empirical probabilities, by gender and ethnicity, of reporting an intended major in SME. A gender × ethnicity interaction is apparent. Among Hispanics, Whites, and Asians, men were more than twice as likely as women to intend SME, with male/female odds ratios of 2.1, 2.6, and 2.8, respectively. However, there was no gender difference among Blacks or American Indians, although the standard errors are quite large for the latter. Consistent with much previous research, Black men were as likely as White men to plan on SME, and, surprising for its magnitude, Black women were twice as likely as White women (33% vs. 20%). Indeed, with the exception of American Indians, all other ethnic groups of women were significantly more likely than White women to enter these colleges intending SME. Asians, within each gender, were the most likely to intend SME.

TABLE 1. Variable Coding and Descriptive Statistics by Gender and Ethnicity for SME-Intending Participants with Complete Data ($N = 5{,}047$)

			Female (43%)					Male (57%)				
Variable	Statistic	Total ($N = 5{,}047$)	Am. Indian ($n = 5$)	Black ($n = 215$)	Hispanic ($n = 61$)	White ($n = 1{,}601$)	Asian ($n = 300$)	Am. Indian ($n = 8$)	Black ($n = 131$)	Hispanic ($n = 81$)	White ($n = 2{,}291$)	Asian ($n = 354$)
SMEgrad	(yes)	55%	0	32%	39%	49%	53%	25%	45%	47%	60%	72%
HSGA	*M*	7.2	6.6	6.6	7.0	7.3	7.4	6.8	6.3	7.1	7.1	7.3
	(*SD*)	(1.0)	(1.1)	(1.1)	(1.0)	(0.9)	(0.8)	(1.6)	(1.3)	(1.0)	(1.0)	(0.9)
	G-M	0.0	−.5	−.5	−.2	.2	.2	−.3	−.9	−.1	0	.1
	(*SD*)	(1.0)	(1.3)	(1.1)	(0.9)	(0.9)	(0.8)	(1.5)	(1.2)	(1.0)	(1.0)	(0.9)
SATM	*M*	649	492	519	584	633	659	636	573	641	671	696
	(*SD*)	(86)	(75)	(82)	(99)	(77)	(78)	(89)	(79)	(93)	(76)	(70)
	G-M	0	−125	−118	−69	−9	4	18	−76	−30	21	25
	(*SD*)	(75)	(66)	(72)	(88)	(67)	(70)	(91)	(75)	(89)	(67)	(58)
SATV	*M*	572	474	499	530	581	577	571	506	569	579	568
	(*SD*)	(93)	(126)	(97)	(93)	(86)	(94)	(94)	(103)	(96)	(88)	(109)
	G-M	0	−64	−73	−48	11	−10	15	−67	−22	8	−13
	(*SD*)	(80)	(126)	(78)	(83)	(73)	(84)	(70)	(92)	(84)	(75)	(97)

Notes. *M* = sample mean, *G-M* = centered around respective institutional means. C&B = *College & Beyond* transcript data. CIRP = Cooperative Institutional Research Program self-report data. URminority = Underrepresented minority = American Indians, Blacks, and Hispanics.

Variable	*Description*, *Coding and* (*Data Source*)
SMEgrad	1 if the student graduate from college with a major in science, math or engineering, 0 otherwise. (C&B)
Gender	Dummy-coded, then grand-mean (.57) centered. Male = .43, Female = −.57. (C&B)
Ethnicity	Contrast-coded, then grand-mean centered. (C&B, primarily, but supplemented by CIRP).
Ethnic-1	White-vs.-URminority: −1.67, .33, −.67, respectively, for URminority, White, Asian; based on −1, 1, 0, then grand-mean (.67) centereing.
Ethnic-2	Asian-vs.-all others: −.39, −.39, 2.61, respectively, for URminority, White, Asian; based on −1, −1, 2, then grand-mean (−.61) centering.
HSGA	Average high school grades, 1–8. 1 = D, 2 = C, 3 = C+, 4 = B−, 5 = B, 6 = B+, 7 = A−, 8 = A or A+. (CIRP)
SAT V/M	College Board *SAT* scores, possible range 200–800 (C&B, except 39 cases imputed from CIRP; see Endnote 8).
INSTSAT	Institutional mean of SATV + SATM for all (96%) of 1989 matriculating students; grand $M = 1220$, $SD = 87$, range = 1046 − 1376.

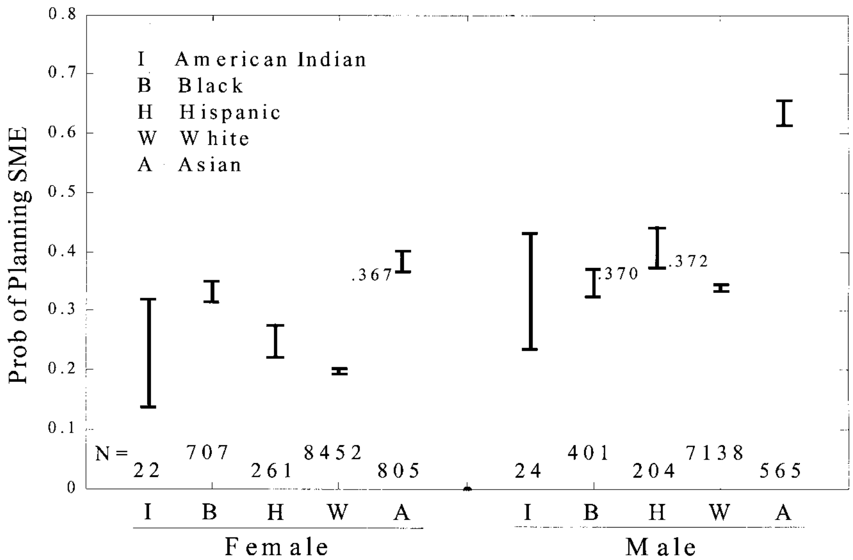

FIG. 2. Observed probability (within ±1 std error) of initially reporting a planned major in SME, by gender and ethnicity for $N = 18{,}579$ students entering 23 C&B/CIRP-participating colleges in Fall 1989.

Science, Mathematics, or Engineering (SME) Graduation as Dependent Variable

The outcome of interest is SME graduation or not (SMEgrad). The coding and descriptive statistics for this and all other variables is included in Table 1. Students were judged SME graduates if they met two criteria: (a) majored in a subject designated as SME, and (b) graduated according to C&B records. Our classification of SME majors follows a consistent research tradition, similar to that outlined by Elliott et al. (1995), including those "traditionally part of natural science divisions: hierarchical, laboratory based disciplines with several prerequisites, usually including many math courses, and usually with heavy workloads and frequent assignments" (p. 7). It excludes social sciences (following, e.g., Astin and Astin, 1993; Hilton et al., 1989), but includes computer science and premedical and dental studies (a complete list is available from the first author). Overall, 55% graduated with a SME degree. Figure 3 is a plot of the empirical probability of SME graduation by gender and ethnicity. With the exception of American Indian women, among whom none of the five SME-intenders persisted, the underrepresented minority students (URminority = American Indians, Blacks, and Hispanics) did not differ significantly from one another in their

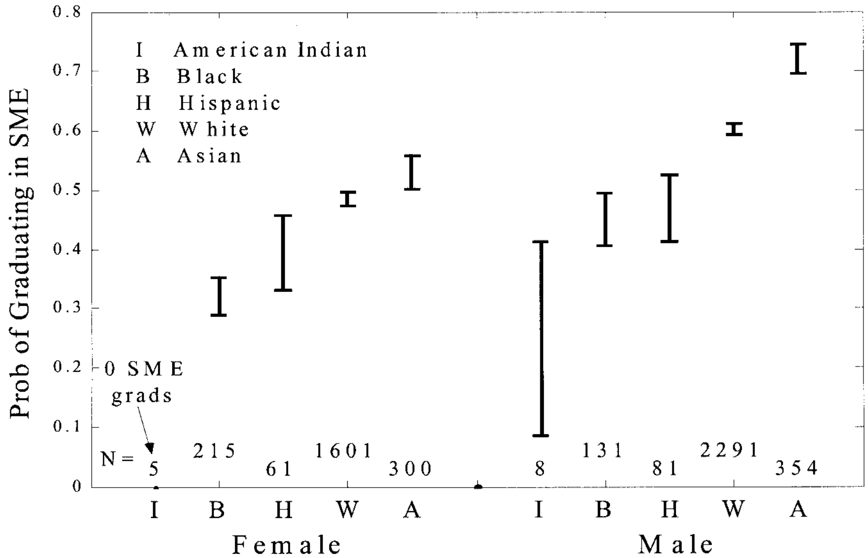

FIG. 3. Observed probability (within ±1 std error) of graduating in SME, by gender and ethnicity, for $N = 5{,}047$ students who initially planned a SME major and had complete data for variables of this study.

observed SME graduation rates. Whites of each gender were more likely than URminorities, and Asians of each gender were more likely than Whites, especially among men. A gender effect favoring males is apparent for each ethnic group except Hispanics.

Student-Level Independent Variables

Ethnicity was derived primarily from college applications (Bowen and Bok, 1998), but supplemented by answers to CIRP ethnicity questions. Following Elliott et al. (1995), Simpson (2001), and our own preliminary analyses,[6] American Indians, Blacks, and Hispanics are grouped together as URminority students. Also like Simpson, but unlike Elliott et al., Asians are not collapsed together with Whites. Simpson concluded

> that European Americans do not significantly differ in their choice of [SME] major from African Americans, Hispanic Americans, and Native Americans. Instead, the significant differences . . . occur between Asian Americans and non-Asians. (p. 88)

Our hypotheses about the differences between these three ethnic groupings, URminorities, Whites, and Asians, were represented by two *a priori* questions and indexed by two orthogonal contrast codes (Cohen and Cohen, 1975, p. 195): is there a significant difference in SME graduation between (a) URminorities and Whites, and (b) between Asians and the other ethnic groups combined (i.e., Whites and URminorities together). Once constructed, each code was grand-mean centered (i.e., around the mean of all 5,047 students in the sample; Table 1) so that the intercept could be interpreted as the expected SME graduation chances for the "average" student. Gender classification was likewise obtained from college applications, and 57% of the participants were men. A dummy code (0 for females, 1 for males) indexes gender and is also grand-mean centered.

High school grade average (HSGA) is a self-report[7] from the CIRP survey. The scale ranged from 1 (D) to 8 (A/A+), with $M = 7.2$, $SD = 1.0$. Not surprisingly amid this selective college sample, this distribution is negatively skewed; less than 8% reported an average below 6 (B+), nearly 50% chose option 8 (A/A+), and more than three-quarters reported A− or higher. In order to directly test our hypothesis about the within-institution relative effect of high school grades, and following the recommendation of Raudenbush and Bryk (2002, p. 143) we centered HSGA around each participant's college mean (i.e., HSGAc = $HSGA_{ij} - HSGA_j$, where ij indexes student i in institution j). Thus, a student with HSGAc = 0 is at the mean among SME-intending students at his/her institution, and nonzero values are deviations from that mean. The within-institution variability (*SD* of HSGAc = 1.0) is the same as for the overall sample (HSGA) variability. While women's average HSGAc is slightly higher than men's ($M = .08$, $SD = .92$ vs. $M = -.06$, $SD = .99$, respectively), this trend is reversed among American Indians and Hispanics. Ethnic differences are apparent within each gender, but are greater among men; for example, Black males averaged about one *SD* lower than Asians and Whites.

Students' SAT math (SATM, $M = 649$, $SD = 86$) and verbal score (SATV, $M = 572$, $SD = 93$), which came primarily from transcripts,[8] were also college-mean centered. Unlike HSGA, the overall sample variability of these scores is greater than the within-institution variability (*SDs* of 75 and 80, respectively, for SATMc and SATVc). Men averaged about 35 points, or just under half a *SD* higher than women on SATMc. This gender disparity was larger among URminorities and smallest among Asians. For SATVc, the direction of gender differences varied by ethnicity; Asian and White women were trivially higher (3 points) than their male counterparts, while males were higher among the URminority samples. For all of our statistical analyses, the centered SAT variables will be scaled by a 75-point increment, so that estimated coefficients may be interpreted as the change in SME graduation probability associated with roughly a 1 *SD* change in the given score (exactly 1 *SD* for the theoretically more critical math score).

College-Level Independent Variable: Selectivity

Institutional selectivity will be indexed by mean institutional SATV + SATM, based on *all* incoming 1989 freshmen with valid SAT scores (i.e., >96%, n = 26,208). This is the same operationalization used by Bowen and Bok (1998), except that they trichotomized the continuous variable, creating three selectivity groups. This variable (INSTSATc) is centered around the grand mean (1220) of the 23 institutions ($SD = 87$, range = 1046–1376) and scaled in analyses by a 90-point increment, also to approximate the effect of a 1 *SD* change.

Analyses

Since our dependent variable is dichotomous, we will model the log-transformed odds of SME graduation so that regression assumptions of linearity and normality are not obviously violated (Hosmer and Lemeshow, 1989; McCullagh and Nelder, 1989). This transformation, denoted η_{ij}, is the natural logarithm of the odds (the ratio of the probability of success, p, to the probability of failure, $1 - p$) for student i in college j. We will fit a progression of HLMs to account for this outcome: [a] first an *unconditional* model, that is, without predictors at either student or institutional levels, will establish a baseline of the between-institution variation (*random-intercept*) in SMEgrad; [b] the fixed effects of student-level variables, constrained to be equal across colleges, will be estimated in a sequence of *random-effects ANCOVA* models, beginning with ethnicity and gender (to estimate the group differences in SMEgrad that are our substantive concern), then including precollege academic variables; [c] a *random-coefficient* model will estimate whether the effects of these student-level indicators vary across institutions; and, finally [d], college selectivity will be included in *intercepts- and slopes-as-outcomes* models to assess whether this alters the fixed estimates or accounts for any cross-college variation in effects indicated by the random-coefficient model. Space limits preclude showing each equation, but our notation follows Raudenbush and Bryk (2002) and, as an example, the level [d] model is written at the student-level (SL) and college-level (CL) in the following form:

$$\text{(SL)}\ \eta_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij}) + \beta_{2j}(X_{2ij}) + \ldots + \beta_{Qj}(X_{Qij}),$$
$$\text{(CL)}\ \beta_{0j} = \gamma_{00} + \gamma_{01}(W_{1j}) + u_{0j} \text{ and } \beta_{Qj} = \gamma_{Q0} + \gamma_{Q1}(W_{1j}) + u_{Qj},$$

where β_{0j} is the average log-odds of SMEgrad for students in college j,

β_{Qj} is the change in log-odds at college j for a 1-unit change in the Qth SL predictor,

X_{Qij} is student is score on the Qth SL predictor,

γ_{00} is the grand mean log-odds of SMEgrad among the colleges,

γ_{Q0} is the average regression slope of predictor Q among the colleges,
$\gamma_{01}(W_{1j})$ is the effect of CL predictor W (= selectivity) on the average log-odds at college j,
$\gamma_{Q1}(W_{1j})$ is the effect of CL predictor W on the effect of SL predictor Q in college j (i.e., the *cross-level interaction*),
u_{0j} is the unique, or unexplained, effect of college j on mean log-odds of SMEgrad,
u_{Qj} is the unique effect of college j on the slope of SL predictor Q.

With 23 colleges as the second-level units, a 5% test level (i.e., $\alpha = .05$) will be highlighted in reports of the HLM estimates. All two-way student-level variable interactions were tested throughout and only statistically significant effects are reported.

RESULTS

Multilevel model results are shown in Table 2. With no predictor variables, model M_0 provides an unconditional estimate of the intercept log-odds of SME graduation, $\gamma_{00} = .14$, and of the statistically significant ($t = 2.8$) cross-college variation in this intercept ($\tau_{00} = .13$). Thus, at the average college in this set, the estimated probability of SME graduation $= \exp\{.14\}/(1 + \exp\{.14\}) = .53$. The intraclass correlation (ρ_I) of .038 indicates only modest clustering or dependency of outcomes at given colleges. Though this model has no predictors, because we assume a fixed value of the level-one residual variance ($\sigma^2_R = 3.29$[9]), it is reasonable to calculate baseline proportions of explained variance against which to compare subsequent models. Following Snijders and Bosker (1999), the total variance (σ^2_Y), that is, the variance of the log-odds of SME graduation, is the sum of student-level variation (3.29) and college-level variation ($\tau_{00} = .13$), or $\sigma^2_Y = 3.42$. Baseline explained variance at the student-level (R^2_1) then $= 1 - (\sigma^2_R/\sigma^2_Y) = .04$, and at the college-level (R^2_2) $= 1 - (\tau_{00}/\sigma^2_Y) = .96$. We do not expect much change in the latter statistic, since the within-college centering of our two student-level academic variables has removed a substantial source of the between-college differences, that is, the means.

Model M_1 estimates SME graduation as a function of ethnicity and gender. The new intercept estimate, .54, is just slightly higher than in M_0. Intercept variation between the colleges (τ_{00}) is reduced from .13 to .10. The variance of the log-odds estimated by this model (σ^2_F) is .11, a statistic that is now added to residual student- and college-level variation to obtain σ^2_Y. Overall, gender and ethnic differences account for 3% of the variance in SME graduation ($R^2_{\text{dicho}} = \sigma^2_F/\sigma^2_Y = .11/3.50 = .03$). Beyond the baseline model, an additional 2% of the variation is explained at the student-level ($R^2_1 = .06$), and 1% at the college-level ($R^2_2 = .97$). Both ethnic contrasts are statistically significant. With gender

TABLE 2. Multilevel Models for Estimating the Log-Odds of College Science, Math, or Engineering Graduation

		Models					
Parameter/Statistic		M_0	M_1	M_2	M_3	M_4	M_5
Fixed Effect (*Level 1*)							
Intercept	γ_{00}	.14 (1.7)	.17 (2.2)	.19 (2.4)	.16 (1.9)	.17 (2.1)	.17 (2.2)
Ethnic-1 (*White vs. URminority*)	γ_{10}		**.30 (6)**	**.22 (4)**	.06 (1.1)	.01 (0.2)	.03 (0.6)
Ethnic-2 (*Asian vs. All Others*)	γ_{20}		**.22 (7)**	**.18 (6)**	**.12 (4)**	**.10 (3)**	**.11 (3)**
Gender (Male .43, Female −.57)	γ_{30}		**.46 (8)**	**.53 (9)**	**.29 (5)**	**.36 (6)**	**.35 (5)**
HS Grade Avg (*HSGAc*)	γ_{40}			**.35 (11)**		**.31 (10)**	**.28 (8)**
*HS Grade Avg*Gender*	γ_{50}			**.18 (2.8)**		**.18 (2.8)**	.13 (1.8)
SATMc	γ_{60}				**.28 (12)**	**.38 (11)**	**.37 (9)**
Fixed Effect (*Level 2*)							
INSTSATc- >Intercept	γ_{01}						.14 (1.6)
INSTSATc- >Ethnic-1	γ_{11}						.06 (1.8)
INSTSATc- >Ethnic-2	γ_{21}						.04 (1.7)
INSTSATc- >Gender	γ_{31}						−.02 (0.7)
INSTSATc- >HSGAc	γ_{41}						−.04 (1.9)
*INSTSATc- >HSGAc*Gender*	γ_{51}						−.06 (1.5)
INSTSATc- >SATMc	γ_{61}						.01 (0.6)

Random Effect							
Intercept Variance	τ_{00}	**.13 (2.8)**	**.10 (2.6)**	**.11 (2.6)**	**.13 (2.7)**	**.13 (2.7)**	**.11 (2.7)**
Goodness-of-Fit							
−2 LogLikelihood	*−2LL*	6850	6728	6590	6568	6461	6448
Δ *−2LL*	*LRT*		122	260	282	389	402
Degrees of freedom	*df*	22	22	22	22	22	22
Intraclass Correlation	ρ_1	.038					
Variance of the linear predictor	σ^2_F		.11	.31	.25	.44	.40
Total Variance of Y (*SMEgrad*)	σ^2_Y	3.42	3.50	3.62	3.67	3.86	3.80
Proportion Explained Variance	R^2_{dicho}		.03	.09	.07	.10	.11
Prop explained at Level 1	R^2_1	.04	.06	.09	.10	.15	.13
Prop explained at Level 2	R^2_2	.96	.97	.97	.96	.97	.97

Notes. All models were fit with SAS PROC NLMIXED with $N = 5{,}047$, and C (colleges) = 23; MLE (ln) parameter estimates are listed with approximate $t = p/se$ (p) in parentheses. Coefficients in bold type are significant at $p < .05$. *Ethnicity* is contrast-coded, then grand-mean centered: *Ethnic-1* = −1.67, .33, 0, for Whites, URminorities, and Asians; *Ethnic-2* = −0.39, −0.39, 2.61. *HSGAc* is self-reported on 8-point scale (1 = D, 8 = A/A+), centered about the mean of each student's college. *SATMc* is student's SAT math score also centered within each college and scaled so that a 1-unit change = 75 points. INSTSATc is the mean institutional SATV+M score centered about the institutional grand mean (1220) and scaled so that a 1-unit change = 90 points. Following Snijders & Bosker (1999), σ^2_F = observed variance of log-odds of SME graduation estimated by the given model, and σ^2_R (Level 1 residual variance) = $\pi^2/3 = 3.29$. $\rho_1 = \tau_{00}/(\tau_{00} + \sigma^2_R)$; $\sigma^2_Y = \sigma^2_F + \tau_{00} + \sigma^2_R$; $R^2_{dicho} = \sigma^2_F/\sigma^2_Y$; $R^2_1 = 1 - (\sigma^2_R/\sigma^2_Y)$; $R^2_2 = 1 - (\tau_{00}/\sigma^2_Y)$.

effect held constant, Whites are more likely than URminorities to graduate in SME ($\gamma_{10} = .30$, $t = 6$), and Asians are more likely than all others ($\gamma_{20} = .22$, $t = 7$). Estimated odds of SME graduation may be calculated for each ethnic group by including in the equation the corresponding value for each contrast code (Table 1). For example, the odds estimated by this model for an URminority woman would be derived as follows:

$$\begin{aligned}
\text{Log-odds } (\eta_{ij}) &= \beta_{0j} + \beta_{1j}(\text{Ethnic1}) + \beta_{2j}(\text{Ethnic2}) + \beta_{3j}(\text{Gender}) \\
&= .17 + 0.30(-1.67) + 0.22(-.39) + 0.46(-.57) \\
&= -0.679 \\
\text{Odds} &= \exp\{-0.679\} = .51
\end{aligned}$$

For either gender, therefore, this model estimates White SME-intenders about 1.8 times as likely as underrepresented minority ones to graduate in SME (odds ratio = .93/.51), and Asians about 2.6 times more likely than URminorities, 1.4 times more than Whites. The gender effect is also significant ($\gamma_{30} = .46$, $t = 8$), estimating that males, regardless of ethnicity, are approximately 1.6 times as likely as females to persist in SME. All ethnicity × gender interactions were tested and found nonsignificant.

The next two models, M_2 and M_3, highlight, respectively, the effect of HSGAc alone (and its significant interaction with gender) and SATMc alone, before effects of both are simultaneously estimated in model M_4. These interim single-academic predictor models are shown because HSGAc and SATMc were found to vary in their relations to the demographic differences that are our substantive focus. Notably, although both reduced the effect of the White vs. URminority contrast, in the SATMc-only model this ethnic difference is no longer statistically significant ($\gamma_{10} = .06$, $t = 1.1$). That is, when differences in SATMc are taken into account, the Whites/URminorities odds ratio is not reliably different from 1.0 or equal odds. When both variables are included in model M_4, each contributes independently to explaining variation in SME graduation and the gender × HSGAc interaction favoring males is unchanged. This model, which we judge the most parsimonious of those tested, accounts for 10% of the total variance of SME graduation and 15% of the student-level variance. The White vs. URminority difference remains nonsignificant ($\gamma_{10} = .01$, $t = 0.2$), and the Asian vs. others difference is attenuated (Fig. 4).[10] Both the Asian/URminorities and the Asian/White odds ratios are now about 1.4. The gender effect depends on level of HSGAc; for example, with HSGAc 1 point below the institutional mean (a standard deviation), men are 1.2 times more likely to graduate, but at 1 point above, they are 1.7 times more likely (Fig. 5).[11] The approximate independent effect of SATM is a 1.5 times, or 50% odds increase in SME graduation per each standard deviation (75 points).

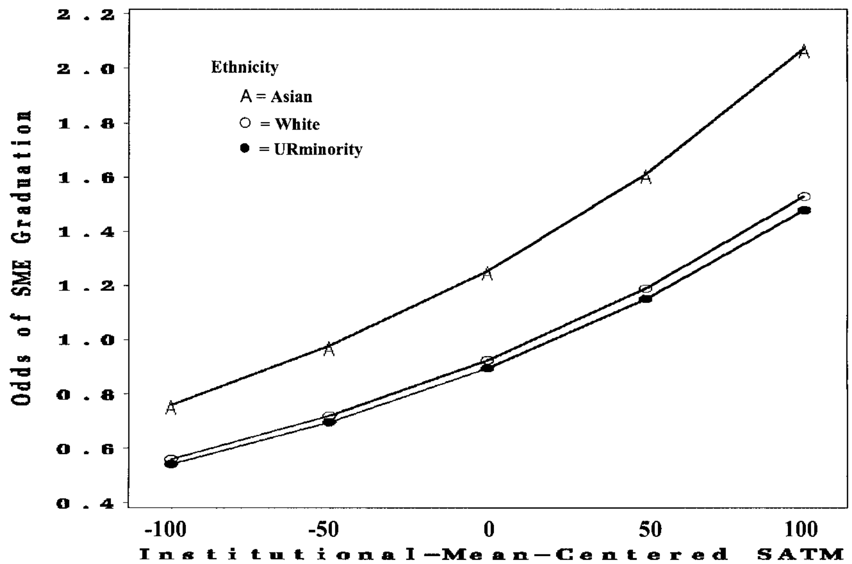

FIG. 4. Model M_4-estimated odds of SME graduation for female students[10] with HSGA = institutional mean, as a function of ethnicity and within-institution-centered SAT math score.

Two additional models were tested at this point, but, since no significant effects were found, their results are not shown in Table 2. First, student-level SAT verbal score had no additional effect. Second, a random-coefficient version of model M_4 was tested and none of the slope variances or covariances were significant. That is, the effects of ethnicity, gender, HSGAc, and SATMc did not vary randomly across institutions and did not covary significantly with the intercept or with one another.

Model M_5, in which we add the indicator of institutional selectivity (INSTSATc), also failed to produce additional significant parameters. Its results are shown in Table 2, however, because this model tests our second hypothesis. Since the random-coefficient model indicated no significant slope variances, INSTSATc was added as a predictor of the random intercepts and of the nonrandomly varying slopes (Raudenbush and Bryk, 2002). In other words, though there was no random slope variation for a college-level variable to explain, it still may be that a college-level variable is systematically related to the fixed effect of a student-level predictor. As indicated in Table 2, INSTSATc did not have a significant effect on any of the parameters, neither on the fixed intercept estimate nor on any of the slope effects. This null result for the estimated effect of

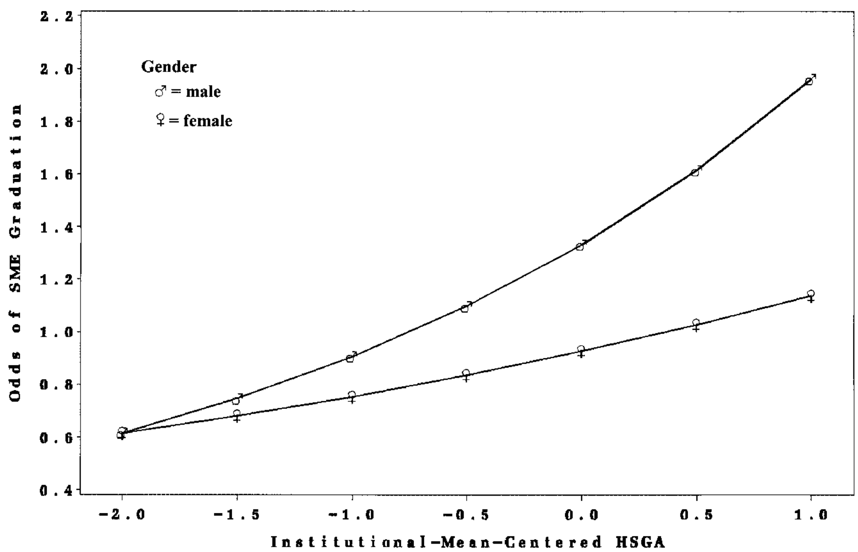

FIG. 5. Model M_4-estimated odds of SME graduation for underrepresented minority students[11] with SATM = institutional mean, as a function of gender and within-institution centered HS grade average.

selectivity on the intercept—that is, on the log-odds of SME graduation for the average student—also obtained when INSTSATc was added only to the equation for the intercept, as opposed to the specification of M_5, in which it is included in the equation for each student-level predictor.

We think it is useful to note here that we fit the equivalent sequence of models (M_0 to M_5) using standard unilevel logit equations, that is, using only the student-level portion of the HLM equation, without the second-level modeling of the β_{Qj} parameters. Thus, we ignored the nested structure of the data in the same manner as Bowen and Bok (1998). For the true student-level variables, this approach yielded log-odds estimates largely similar to those of the multilevel models. However, the estimated main effect of INSTSATc, treated now as if a score for each student, was significant at $p < .001$. Though the estimated point effect of INSTSATc is essentially the same under both approaches (increased likelihood of SME graduation for each 90-point increase in INSTSATc estimated at 1.13- and 1.15-times, respectively, with unilevel and multilevel models), because the unilevel approach disaggregates the college variable to the student level, the standard errors are calculated based on an $N > 5{,}000$, rather than $N = 23$.

DISCUSSION

We have presented statistical models of SME graduation for $N = 5{,}047$ American Indian, Asian, Black, Hispanic, and non-Hispanic White students who reported an intention to major in SME when they began at 23 selective colleges in the fall of 1989. Two hypotheses were tested: [1] that among students initially intending a SME major, differences in precollege academic preparation will account for ethnic and gender differences in SME graduation rates; and [2] that college selectivity, beyond effects of student characteristics, will be positively associated with SME graduation. The colleges in the sample are mostly highly selective, so generalization of our findings to the population of all college matriculants is unwarranted. Within the limits of this sample, however, the multilevel analytic approach we used accounts for the biases that would otherwise result from the nonrandom grouping of students at these colleges.

Summary Results

Empirical SME graduation rates for Asians, Whites, and underrepresented minorities were 63%, 55%, and 38%, respectively, 61% for men, 47% for women. Hypothesis 1 was fully supported with respect to the SME graduation difference between White and underrepresented minority students, and partially supported for that between Asian and all other students, and between genders. Specifically, with gender differences accounted for, Whites were 80%, or 1.8 times more likely than underrepresented minorities to graduate in SME. But taking into account students' scores on SATM relative to the other SME-intenders at their college completely explained this differential, that is, reduced the estimated difference in likelihood effectively to zero. Differences in relative HSGA did not substantively alter this result. Underrepresented minority SME-intenders averaged 85 points below the mean SATM for intenders at their institutions, while White intenders averaged 8 points above (*SDs* were 85 and 69, respectively). Asians, also after accounting for effects of different gender distributions, were 2.6 times more likely than underrepresented minorities and 1.4 times more likely than Whites to persist in SME. They remained statistically more likely even after accounting for relative SATM and HSGA, but their differential over underrepresented minorities was substantially reduced, now estimated at 1.4 times greater odds (while the Asian vs. White difference remained at approximately 1.4, indicating that this gap is not related to differences on HSGA or SATM). Accounting for precollege academics also reduced the observed male/female odds ratio of 1.6 (with ethnicity held constant), but the magnitude of this effect varied for students at different levels of HSGA. With relative SATM differences accounted for, the estimated male advantage was reduced to about 1.4 times if HSGA was at the institutional mean, was still lower (1.2) for HSGA 1 point below the mean, but was higher (1.7) for HSGA 1 point above the mean.

Overall, with other student-level variables held constant, the estimated effect of a 75-point (= 1 *SD*) increase in a student's relative SATM is a 50% or 1.5 times improvement in SME-graduation odds. This effect functioned similarly for men and women and for underrepresented minorities, Whites, and Asians. That is, results of tests of interaction effects of SATM with both ethnicity and gender were nonsignificant. With relative SATM held constant, a 1-point increase in a student's relative high school grades (= 1 *SD*) was associated with a 40% or 1.4 times improvement in odds of SME graduation, though we have noted that this effect varies with gender. Considering students' SAT verbal scores did not improve on the prediction of SME graduation once HSGA and SATM were accounted for.

Our multilevel modeling approach allowed tests of whether the effects of these student-level predictors on SME graduation vary randomly across institutions. We found no significant random variation of effects across this set of colleges. This did not rule out, however, systematic variation of effects across colleges, when the student-level slope effects "vary strictly as a function of [a college-level variable] W_j" (Raudenbush and Bryk, 2002, p. 28). Thus, we proceeded to test the college-level hypothesis (Hypothesis 2) that students of comparable demographic status and high school academic credentials will be more likely to graduate in SME at more selective colleges. Support was not found for this hypothesis. With ethnicity, gender, relative high school grades and SATM scores accounted for, the effect of selectivity was nonsignificant; it influenced neither the average odds of SME graduation nor the functioning of student-level predictors, that is, cross-level interactions.

We replicated our analyses using standard unilevel logit regression and treated the college-level variables as if they were scores at the student-level. This was the approach used in all previous multicollege studies of SME persistence that we reviewed, and also by Bowen and Bok (1998) with the C&B data, but for different outcomes. Despite our multilevel finding that relatively little of the variance in students' SME persistence was at all related to college differences, employing a unilevel (student-level-only) approach—owing to misestimated standard errors for the college-level effects—resulted in finding a significant positive effect of institutional selectivity. Thus, had we not accounted for students' "nesting" within colleges by using the multilevel approach, we would have arrived at a spurious conclusion about the effect of institutional selectivity.

Interpretation and Implications

This study was not designed to provide a comprehensive model of SME persistence at selective colleges. Rather, it applied conceptually simple models using the most basic, commonly collected precollege academic information to answer questions about substantial ethnic and gender differences in SME graduation. Our best model (M_4) is based on just two student-level predictors beyond ethnicity and gender—relative high school grades, and relative SAT math score—and accounts for a modest 10% of the variance in SME graduation. This simple model, however (indeed, an even simpler one without HSGA), was sufficient to explain the observed 80% greater likelihood of SME graduation of Whites compared with underrepresented minorities who began at these 23 colleges with the expressed intent to major in SME. It was also sufficient to reduce Asians' empirical 160% greater chances compared with underrepresented minority students to 40% greater, and reduce men's observed 60% greater chances compared with women to, on average, 40%.

Bowen and Bok (1998) explained the "very large" difference in mean collegiate class rank between 1989 cohort *College & Beyond* Blacks and Whites (23rd vs. 53rd percentile, respectively) as follows:

> A student with a given SAT score, high school grades, and so on, who attends one of the most selective schools, should be expected to have a lower rank in class than a student with the same credentials who attends a school that enrolled a smaller number of top-rated students. This is precisely the pattern we found. (p. 73)

Now we have found a comparable pattern with respect to persistence in science, math, or engineering majors. In this case, however, the cost to a group with systematically lower relative preparation is a greater likelihood of being on the negative end of a "yes or no" outcome, rather than being lower on a continuum. Elliott and his colleagues (1995) concluded that race-sensitive admission, while increasing access to elite colleges, was inadvertently causing disproportionate loss of talented underrepresented minority students from science majors. Our findings for the *College & Beyond* students are consistent with this inference. According to our model (and assuming that our sample represents all 1989 matriculants at these 23 colleges), if all of the SME-intending underrepresented minority students had enrolled in similarly functioning colleges where their high school grades and math test scores averaged at the institutional means among SME intenders, 72 more of the women and 62 more of the men would be predicted to persist in SME (45% and 35% increases, respectively). This projection is based both on untested and untestable assumptions, for example, that the model applies beyond this selectivity range of colleges and, more fundamentally, that these correlational findings would turn out the same under conditions of random assignment. But to the extent that this inference is correct and holds for colleges nationwide of similar selectivity and across cohorts other than the collegiate class of 1993, we agree with Elliott et al. that the implications for many talented minority students warrant serious consideration.

Our finding that the effect of relative SAT math score functioned in the same way for all ethnic groups, across gender and across colleges, has implications for those choosing colleges, for those choosing students, and for efforts to improve precollege math and science preparation for all students. Finding, like us, that "ethnic variables added a zero amount" to the prediction of SME graduation after accounting for precollege academic preparation, Hilton et al. (1989) called for efforts to close quantitative ability gaps prior to college:

> In selecting students to be majors in SME fields, there are no grounds for considering the applicant's ethnicity if the applicant is otherwise qualified. . . . This is not to say that efforts to assist minority students in their SME careers should be abandoned. Rather it is to say that: A primary effort of secondary school science instruction should be to help interested minority students to become qualified for SME majors in college. (p. 168)

Chipman and Thomas (1987) found that ethnic differences in math achievement are already apparent among young school children and concluded that attempts to close the math gap "must give high priority to the improvement of school achievement from the earliest grades" (p. 414). By the end of the high school years, national differences on admission test scores across ethnic groups and gender reflect a myriad of unequal social and economic conditions that bear on educational opportunity and aspiration. But among the students enrolling at these selective colleges, substantial test score differences between demographic groups can be attributed primarily to admission and enrollment decisions. To the extent that SAT math scores of different groups of admitted SME-intending students systematically vary from one another within the same college, our analysis suggests that systematic group differences in SME graduation rates can be expected. This should not be interpreted to mean that when choosing between two science-interested students, the one with the higher math SAT *always* ought to be selected; admission officers know well that "other factors" usually are not "equal." Evidence of particularly strong intrinsic motivation to study science, for instance, may suggest that the lower-scoring student is more likely to succeed. The statistical findings indicate, rather, that relatively higher math SAT scores can be expected, on average, to be associated with higher likelihood of science persistence, regardless of ethnicity or gender.

Our model indicates that the same basic relation holds for the effect of high school grades, though college admission decisionmakers must try to compare grades that are based in many different metrics. We found that accounting for high school grade differences alone was not sufficient to explain the SME graduation gap between Whites and underrepresented minorities, probably because, as Culotta and Gibbons (1992) point out, all high school math and science courses of the same name are not of the same rigor. A possible explanation for our finding that the same increase in relative high school grades is associated with a greater increase in estimated prospects of SME graduation for men than women is that boys are more likely than girls to take advanced math and science courses (Chipman and Thomas, 1987; Farmer, Wardrop, Anderson, and Risinger, 1995; Jacobs and Wigfield, 1989; Lee, 1987; Leslie et al., 1998; NAS, 1987; Oakes, 1990). Thus, to the extent that higher grade-averages for boys are disproportionately comprised of more challenging SME-relevant courses, their grades are not "the same," and it makes sense that they would fare better in college math and science.

This course selection difference is also among the many possible contributors to the remaining unexplained gender effect in this study. Other potentially relevant factors include differences in self-efficacy and confidence in SME domains, and different ratings of the usefulness of SME (Betz and Hackett, 1983; DeBoer, 1986; Leslie et al., 1998; Strenta et al., 1994), different reasons for intending a SME major, e.g., more or less intrinsically motivated, or thing- vs. people-interested (Leslie et al., 1998; Seymour and Hewitt, 1997; Strenta et al., 1994), and unconscious or "implicit" stereotypes (Nosek, Banaji, and Greenwald, 2002; Spencer, Steele, and Quinn, 1999). At the elite colleges studied by Strenta et al., men were somewhat more likely than women to do unassigned reading related to their SME classes, and women were more likely than men to judge the classes as too competitive, though they did not report evidence of a "chilly" or discriminatory climate for women.

Some of the same factors, as well as others, may account for the remaining advantage of Asian students. Chipman and Thomas (1987) noted that Asians were significantly more likely to take trigonometry, calculus, chemistry, and physics. The Asian SME-intenders studied by Elliott et al. (1995) took substantially more high school science courses than any other group. Seymour and Hewitt (1997) found that an extrinsic basis for choosing SME—family expectation—was particularly influential for Asians. Fullilove and Treisman (1990) found that Chinese Americans at U.C. Berkeley were more likely to combine their studying with socializing, that is, by studying in groups.

After accounting for the effects of student characteristics, there was no additional improvement in SME chances associated with attending a more selective college. We illustrated that failure to employ a multilevel statistical analysis would have resulted in a different answer about the effect of institutional selectivity. If the variables needed for our analysis were available for all 28 of the colleges analyzed by Bowen and Bok (1998), it is possible that characteristics of the additional five schools and the increased statistical power may have resulted in a significant estimated effect of selectivity. Still, if our models did not take into account the within-institution dependencies among students selected, and selecting, into these colleges, we would not know if the answer was correct. As Snijders and Bosker (1999) concluded, "if the macro-units have any meaningful relation with the phenomenon under study, analyzing only aggregated or only disaggregated data is apt to lead to misleading and erroneous conclusions" (p. 16).

The National Science Board (2002) reiterated recently that fostering minority and female interest in science majors should be a national priority. Our preliminary descriptive analysis of gender and ethnic differences in the likelihood of planning a SME major at the point of entry to these colleges suggests that underrepresented minority students may not have been less interested than Whites. Indeed, among women, Blacks and Hispanics were significantly more likely than Whites to declare a SME major. Gender differences favoring men, however, were substantial among Asians, Hispanics, and Whites. Much research has been devoted to understanding these precollege-formed gender differences in the development of science interest, but their ongoing manifestation among students enrolling at highly selective colleges should motivate continued scrutiny.

The focus of our analysis, however, has been to clarify to what extent information readily available to admission officials can account for the disproportionate attrition from science majors of the initially interested underrepresented minority and female students at these selective colleges. Our findings suggest that if the standardized mathematics test scores or high school grades of any group of enrolling students—regardless of ethnicity or gender—are systematically lower than others', then the SME graduation rate of that group is also likely to be lower. These often-replicated findings suggest a conundrum for selective college admission officials. At the individual level, offering a relatively educationally disadvantaged applicant the chance to benefit and graduate from a more selective institution may put at increased risk his or her goal of a career in science. The same potential trade-off is suggested for group-level equity goals; faster increase in representation at elite colleges, but slower in scientific fields vs. slower increase at elite colleges, but faster in scientific fields. Each of these sets of individual and group equity goals is widely agreed to be of critical importance, and how best to achieve them ought to be a matter of well-informed discussion and further national inquiry.

ACKNOWLEDGMENTS

Versions of this research have been presented at Annual Meetings of the American Educational Research Association (2002) and The College Board (2003). We are grateful to the officers of the Andrew W. Mellon Foundation for allowing access to the *College & Beyond* data and for their constructive questions and suggestions about the design of our study. Other research based on the *College & Beyond* data includes Bowen and Bok's *The Shape of the River* (1998) and Shulman and Bowen's *The Game of Life* (1998). We appreciate the suggestions on preliminary drafts provided by Rogers Elliott, Suzanne Estler, John Nesselroade, John Smart, and two anonymous reviewers. We also thank Fumiaki Hamagami for generously providing technical insights on multilevel analytic and graphical programming.

ENDNOTES

1. Less than 2% of students entering in 1989 had a reported ethnic status of "Foreign," "Unknown," or "Other." These cases were not included in our analyses.
2. Twenty-four colleges participated in the CIRP, but one was excluded because none of its $N >$ 1,500 students had a response for the question about high school grade average. The 23 institutions included are Barnard, Bryn Mawr, Columbia, Denver, Emory, Hamilton, Kenyon, Miami of Ohio, Northwestern, Oberlin, U. Penn, Princeton, Penn State, Rice, Smith, Swarthmore, Tufts, Tulane, UNC Chapel Hill, Vanderbilt, Wellesley, Wesleyan, and Williams.
3. The most obvious potential source of selection bias in this study is the 28% of students at CIRP-participating colleges from whom a CIRP survey was not obtained. After this source, few participants had missing values on our variables of interest: 95% had a valid report of intended major, 97% a valid SAT score, and 99% had valid HS grade average and SME graduation status. We assessed the degree to which missing a CIRP survey is related to variables of interest in this study by logit-regressing survey completion (yes or no) on (a) ethnicity and gender (n = 26,665), and (b) on ethnicity, gender, SAT math score, and SME graduation status (n = 25,258). Each model accounted for only 1% of the variance in CIRP participation. It was also found that adding CIRP participation to a logit model with ethnicity, gender, SAT math and verbal score as predictors of SME graduation (our study DV) made no improvement in the prediction (n = 25,258; HS grade average could not be used, as its source was the CIRP survey). Given these small relations between CIRP participation and variables of theoretical interest, additional statistical techniques to account for any biases will prove ineffective. Consequently, we will assume that our sample of SME-intenders is representative of the SME intenders at the 23 CIRP colleges and that data are "missing completely at random" (Little and Rubin, 1987).
4. All reported N and analyses are weighted to account for the C&B sampling design at two of the institutions (Bowen and Bok, 1998).
5. The CIRP questionnaire contained a list of 81 undergraduate majors (including "undecided") and instructions to "Mark only *one* oval to indicate your probable field of study" (Higher Education Research Institute, 1989, p. 4). For purposes of this study, responses were dichotomized as either SME or not-SME.
6. Initially, we excluded the 13 American Indians from the analysis and constructed three orthogonal contrast codes to represent the remaining four ethnic groups: one code contrasted Blacks vs. Hispanics, one Whites vs. Blacks/Hispanics together, and one Asians vs. Blacks/Hispanics/Whites together. When the Black vs. Hispanic code was nonsignificant in the prediction of SMEgrad, we joined Blacks and Hispanics in the URminority group and included the American Indians.
7. Our confidence in the accuracy of these reports is high, because self-reports of SAT scores on the same CIRP survey were very accurate; official and self-reported scores were available for $n > 14{,}000$ C&B/CIRP respondents, and Pearson's r for the verbal and math scores, respectively, was .92 and .93.
8. For all but 39 of the $N = 5{,}047$ complete-data participants, the source for SATM was C&B data from college transcripts. The other 39 have a score imputed from their self-report on the CIRP survey (see note 6). The obtained regression equation was: Transcript SATM = 45.3 + .9282 (Self-report SATM).
9. Because the student-level residuals are limited to one of two values ($-P_j$ if SMEgrad = 0, or $1 - P_j$ if SMEgrad = 1; Raudenbush and Bryk, 2002; Snijders and Bosker, 1999), we assume that the level-one residuals have a standard logistic distribution with a mean of zero and variance (σ^2_R) of $\pi^2/3 = 3.29$ (Snijders and Bosker, 1999, chap. 14).
10. Odds estimates in Fig. 4 were calculated as if these students were female, an arbitrary choice because the effects of interest in this plot, ethnicity and SATM, do not depend on gender.

11. Odds estimates in Fig. 5 were calculated as if these students were URminorities, an arbitrary choice because the effects of interest in this plot, gender and HSGA, do not depend on ethnicity.

REFERENCES

Adair, R. K. (1991). Using quantitative measures to predict persistence in the natural sciences. *College and University* **Fall:** 73–79.

Astin, A. W., and Astin, H. S. (1993). *Undergraduate Science Education: The Impact of Different College Environments on the Educational Pipeline in the Sciences*, Higher Education Research Institute, UCLA, Los Angeles.

Barron's Educational Series, Inc. (2001). *Profiles of American Colleges*, Barron's Educational Series, Inc., Hauppauge, NY.

Betz, N. E., and Hackett, J. (1983). The relationship of self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior* **23:** 329–345.

Bowen, W. G., and Bok, D. (1998). *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*, Princeton University Press, Princeton, NJ.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. In: Berliner, D. C. (ed.), *Review of Research in Education* **(Vol. 8)**, American Education Research Association, Washington, DC, pp. 158–233.

Buunk, B. P., and Ybema, J. F. (1997). Social comparisons and occupational stress: The identification-contrast model. In: Buunk, B. P., and Gibbons, F. X. (eds.), *Health, Coping, and Well-Being: Perspectives from Social Comparison Theory*, Erlbaum, Mahwah, NJ, pp. 359–388.

Chipman, S. F., and Thomas, V. G. (1987). The participation of women and minorities in mathematical, scientific, and technical fields. *Review of Research in Higher Education* **14:** 387–430.

Civian, J., and Schley, S. (1996, April). *Pathways for women in the sciences II: Retention in math and science at the college level*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Cohen, J., and Cohen, P. (1975). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Culotta, E., and Gibbons, A. (eds.). (1992). Minorities in science: Two generations of struggle [Special Report]. *Science* **258:** 1176–1232.

Davis, J. A. (1966). The campus as a frogpond: An application of the theory of relative deprivation to career decisions of college men. *American Journal of Sociology* **72:** 17–31.

DeBoer, G. E. (1986). Perceived science ability as a factor in the course selections of men and women in college. *Journal of Research in Science Teaching* **23:** 343–352.

Dunteman, G. H., Wisenbaker, J., and Taylor, M. E. (1979). *Race and Sex Differences in College Science Program Participation* (Report No. RT1-22U-1570), Research Triangle Institute, Research Triangle Park, NC. (ERIC Document Reproduction Service No. ED 199 034)

Educational Testing Service (1989). *1989 National SAT Profiles*, Educational Testing Service, Princeton, NJ.

Elliott, R., Strenta, A. C., Adair, R., Matier, M., and Scott, J. (1995). *Non-Asian minority students in the science pipeline at highly selective institutions*. (Final report of NSF Grant RED 93 53 821).

Ethington, C. A. (1997). A hierarchical linear modeling approach to studying college effects. In: Smart, J. C. (ed.), *Higher Education: Handbook of Theory and Research (vol. 12)*, Agathon, New York, pp. 165–194.

Farmer, H. S., Wardrop, J. L., Anderson, M. Z., and Risinger, R. (1995). Women's career choices: Focus on science, math, and technology careers. *Journal of Counseling Psychology* **42:** 155–170.

Fullilove, R. E., and Treisman, P. U. (1990). Mathematics achievement among African American undergraduates at the University of California, Berkeley: An evaluation of the mathematics workshop program. *Journal of Negro Education* **59:** 463–478.

Green, K. C. (1989). A profile of undergraduates in the sciences. *American Scientist* **77:** 475–480.

Higher Education Research Institute (1989). *Cooperative Institutional Research Program (CIRP) Freshman Questionnaire, UCLA*, Higher Education Research Institute, Los Angeles.

Hilton, T. L., Hsia, J., Solorzano, D. G., and Benton, N. L. (1989). *Persistence in Science of High Ability Minority Students*, Educational Testing Service, Princeton, NJ.

Hosmer, D. W., and Lemeshow, S. (1989). *Applied Logistic Regression*, John Wiley, New York.

Jackson, L. A., Gardner, P. D., and Sullivan, L. A. (1993). Engineering persistence: Past, present, and future factors and gender differences. *Higher Education* **26:** 227–246.

Jacobs, J. E., and Wigfield, A. (1989). Sex equity in mathematics and science education: Research-policy links. *Educational Psychology Review* **1:** 39–56.

Jagacinski, C. M., LeBold, W. K., and Salvendy, G. (1988). Gender differences in persistence in computer-related fields. *Journal of Educational Computing Research* **4:** 185–202.

Kelley, H. H. (1952). Two functions of reference groups. In: Swanson, G. E., Newcomb, T. M., and Hartley, E. (eds.), *Readings in Social Psychology*, Henry Holt, New York, pp. 410–414.

Lee, V. E. (1987). *Identifying Potential Scientists and Engineers: An Analysis of the High School-College Transition*, Office of Technology Assessment, Washington, DC.

Leslie, L., and Oaxaca, R. (1998). Women and minorities in higher education. In: Smart, J. C. (ed.), *Higher Education: Handbook of Theory and Research* **(vol. 13)**, Agathon, New York, pp. 304–352.

Leslie, L., McClure, G. T., and Oaxaca, R. (1998). Women and minorities in science and engineering: A life sequence analysis. *Journal of Higher Education* **69:** 239–276.

Levin, J., and Wyckoff, J. H. (1995). Predictors of persistence and success in an engineering program. *NACADA Journal* **15:** 15–21.

Little, R. J. A., and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, John Wiley & Sons, New York.

Marsh, H. W., Kong, C., and Hau, K. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology* **78:** 337–349.

McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models* (2nd ed.), Chapman and Hall, New York.

Meece, J. L., Parsons, J. E., Kaczala, C. M., Goff, S. R., and Futterman, R. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin* **91:** 324–348.

Nakao, K., and Treas, J. (1990). *Computing 1989 Occupational Prestige Scores* (General

Social Survey Methodological Report no. 70), National Opinion Research Center, Chicago.

National Academy of Sciences (1987). *Nurturing Science and Engineering Talent: A Discussion Paper*, The Government-University-Industry Research Roundtable, Washington, DC.

National Science Board (2002). *Science and Engineering Indicators—2002* (NSB 02-01) [On-line], National Science Foundation, Arlington, VA. Available: http://www.nsf.gov/sbe/srs/seind02/start.htm

National Science Foundation (1992). *Blacks in Undergraduate Science and Engineering Education*, National Science Foundation, Washington DC.

National Science Foundation (1999). *Women, Minorities, and Persons with Disabilities in Science and Engineering* (NSF Report No. 99-338), National Science Foundation, Arlington, VA.

Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology* **83:** 44–59.

Oakes, J. (1990). *Lost Talent: The Underparticipation of Women, Minorities, and Disabled Persons in Science*, Rand, Santa Monica, CA.

Pascarella, E. T., and Terenzini, P. T. (1991). *How College Affects Students*, Jossey-Bass, San Francisco.

Post, P., Stewart, M. A., and Smith, P. L. (1991). Self-efficacy, interest, and consideration of math/science and non-math/science occupations among black freshmen. *Journal of Vocational Behavior* **38:** 179–186.

Ramist, L., Lewis, C., and McCamley-Jenkins, L. (1994). *Student Group Differences in Predicting College Grades: Sex, Language, and Ethnic Groups* (College Board Report No. 93-1), College Entrance Examination Board, New York.

Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition*, Sage, Thousand Oaks, CA.

Schoenberger, A. K. (1988, April). *College women's persistence in engineering and physical science: A further study*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Schulman, J. L., and Bowen, W. G. (2001). *The game of life: College sports and educational values.* Princeton, NJ: Princeton University Press.

Seymour, E., and Hewitt, N. M. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences*, Westview Press, Boulder, CO.

Simpson, J. C. (2001). Segregated by subject: racial differences in the factors influencing academic major between European Americans, Asian Americans, and African, Hispanic, and Native Americans. *Journal of Higher Education* **72:** 63–100.

Snijders, R. A. B., and Bosker, R. J. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage Publications Ltd., London.

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* **35:** 4–28.

Strenta, A. C., Elliott, R., Adair, R., Matier, M., and Scott, J. (1994). Choosing and leaving science in highly selective institutions. *Research in Higher Education* **35:** 513–547.

Tobias, S. (1990). *They're Not Dumb, They're Different: Stalking the Second Tier*. Research Corporation, Tucson, AZ.

Toutkoushian, R. K., and Smart, J. C. (2001). Do institutional characteristics affect student gains from college? *Review of Higher Education* **25:** 39–61.

Turner, S. E., and Bowen, W. G. (1999). Choice of major: The changing (unchanging) gender gap. *Industrial and Labor Relations Review* **52:** 289–313.

Ware, N., and Lee, V. (1988). Sex differences in choice of college science majors. *American Educational Research Journal* **25:** 593–614.
Ware, N. C., Steckler, N. A., and Leserman, J. (1985). Undergraduate women: Who chooses a science major? *Journal of Higher Education* **56:** 73–84.

Received July 22, 2002.