Formal and Social-Psychological Elements of Secondary Schooling on College Major Selection

Brian P. An
University of Iowa

ABSTRACT

This paper examines the relation between formal and social-psychological elements of secondary schooling and college major selection. Students select their college major based on a "secondary sorting process" where their high school experiences are key factors that transmit certain advantages for some students and disadvantages for others. Expectations of working in a STEM occupation at age 30 are an essential catalyst for students to select a STEM major. In addition, self-efficacy in math and science influences the likelihood that a student selects a STEM major. Occupational expectations, self-efficacy, and effort completely mediate the relation between parental education and selecting a STEM major. Subject-specific indicators predicted major selection, but precollege math coursework exerted a more substantial influence in choosing a technical–quantitative major than precollege science coursework.

Keywords: college major, STEM, field of study, self-efficacy, occupational expectations

INTRODUCTION

The economic returns for individuals with a bachelor's degree are greater than those without a degree, and this disparity has increased over time (Behrman et al., 1996; Karen & Dougherty, 2005; Li et al., 2012). It is becoming more difficult for high school graduates to procure well-paying manufacturing jobs to sustain a good standard of living. In the United States, Tamborini, Kim, and Sakamoto (2015) estimate the lifetime earnings premium of a bachelor's degree over a high school degree is 43% for men and 51% for women. Nevertheless, the lifetime earnings premium of a bachelor's degree of women is 70% less than that of men ($587,000 and $840,000, respectively).

Although people who obtain a bachelor's degree are often viewed as the "cream of the crop," they are still differentiated by college majors. Social scientists have demonstrated that the discipline students pursue in college is important for their future well-being (Davies & Guppy, 1997; Kim et al., 2015; Webber, 2016). While the relationship is not entirely deterministic, individuals graduating with business and science-oriented degrees have higher lifetime earnings than those from other fields (Kim et al., 2015).

Often termed horizontal stratification, this paper concentrates on stratification within an education system. This contrasts with vertical stratification, which refers to systematic differentiation in the degree or quantity of education received (Charles & Bradley, 2002; Gerber & Cheung, 2008). Specifically, this paper considers formal (e.g., coursework) and social-psychological (e.g., self-efficacy and academic effort) elements of secondary schooling and examines whether these elements act as mechanisms that allocate students into undergraduate fields. In the United States, students choose their college major with fewer restrictions and less rigidity than those from other countries (Duta et al., 2018; Kerckhoff, 2001). Students select their college major partly based on a secondary sorting process. Their high school experiences serve as key influences that transmit certain advantages for some students and disadvantages for others.

This paper differs from previous research in two ways. First, past research on horizontal stratification generally concentrates on how qualitative differences at secondary schools, such as course-taking patterns and school tracks, influence college destinations (Duta et al., 2018; Posselt et al., 2012; Stearns et al., 2013). By contrast, this paper focuses on whether horizontal differentiation at postsecondary schools is affected by horizontal differentiation at a prior transition. This process is "secondary sorting" because the stratified educational experiences in high school impact a student's within-transition (e.g., college majors) outcomes, above and beyond its primary influence on between-level transitions (e.g., college attendance after high school).

Second, unlike previous studies that measure high school course history as the number of subject-specific credits earned, this paper focuses on the academic rigor of high school courses. However, this study does not use Adelman's (2006) influential measure of academic curriculum intensity because it comprises a host of courses in different subject areas. Instead, the interest lies in the link between the rigor of subject-specific courses and college majors. Therefore, indicators that capture a student's academic pipeline along specific content areas are most appropriate.

LITERATURE REVIEW

Precollege Academic Preparation on College Participation

The degree to which college entrants are academically prepared is crucial for postsecondary success. Prior studies generally show that academic preparation in high school significantly impacts college remediation and persistence (Adelman, 2006; An, 2013a, 2013b; Kirst & Bracco, 2004; Parsad & Lewis, 2003). Academic preparation is also associated with college major choice, where students are inclined to select majors in which they had prior exposure to the subject matter (Simpson, 2001). However, previous studies generally concentrate on the number of credits earned rather than the academic rigor of the courses.

By completing coursework, students receive credentials to bargain with when applying to higher education (Duta et al., 2018; Stearns et al., 2013). Hence, high schools sort and select students by typically funneling "qualified" candidates into four-year institutions and channeling "less qualified" candidates out of academics or into two-year institutions where the chance of ultimately completing a four-year degree is lower. Researchers find placement in the high school curriculum has consequences for students' achievement, postsecondary plans, and attainment (Adelman, 2006; Duta et al., 2018; Lucas, 2001; Posselt et al., 2012).

Advanced high school math courses are also crucial in individuals' success in college math courses. Students taking first-year undergraduate calculus courses without prior exposure to calculus are in the greatest danger of receiving low marks and being prone to fail the course (Burton, 1989; Ferrini-Mundy & Gaudard, 1992). By contrast, students who take a preparatory year of calculus in high school, while not qualified to "test out" of the material (received Bs and Cs primarily), as a group, receive the highest grades in the course.

Precollege influences help explain the maintenance or persistence of societal realities. For example, non-merit factors, such as socioeconomic and race/ethnicity, play a role in student track placement and course-taking patterns. Studies show that racially minoritized and low-income students are more likely located in lower secondary school tracks than White students (Lucas, 1999; Tyson, 2011). Research further shows that parents contribute to the course placement of their children. Middle-class parents, perhaps with more extensive knowledge of the importance of courses on their children's education, are more involved with their children's well-being and, if necessary, dismiss and overrule teachers' recommendations by insisting on their child's placement in advanced courses (Lareau, 2011; Lucas, 2017; Useem, 1992).

At the high school level, gender differences exist in course placement. In eleventh and twelfth grade, girls are more likely to combine non-college prep math and college-prep English than boys (Lucas, 1999). This difference has significant consequences because advanced-level math and science courses in high school are strong predictors in the likelihood of selecting a math or science-oriented college major (Engberg & Wolniak, 2013; Riegle-Crumb et al., 2012; Russell & Atwater, 2005). Therefore, what matters is not simply taking courses piece-meal but instead taking advanced course sequences. This sequence is essential since these courses are still selective, and many students do not take these course sequences—and those who do maintain a premium (Schiller & Hunt, 2011; Stevenson et al., 1994).

Self-Perceptions of Ability and Academic Interest on Educational Outcomes

In addition to course work, academic trajectories are also influenced by social-psychological factors such as academic interest and self-perceptions of ability (Wang, 2013).

Studies show that twelfth-grade math achievement is directly and indirectly affected by tenth-grade math interest (Köller et al., 2001). In college, researchers find that interest continues to impact performance, course selection, and choice of major (Harackiewicz et al., 2000; Harackiewicz et al., 2002). Interest is an interactive engagement between an individual and some aspect of his/her/their environment (e.g., events and ideas) (Hidi & Harackiewicz, 2000). Therefore, interest is content specific and not a predisposition that transcends across all activities.

Scholars have distinguished between two types of interest—individual interest and situational interest. Individual interest refers to a personal disposition for a given domain or task that develops over time, which in turn may lead to increased knowledge and value for that task. Individual interest also refers to the immediate psychological state of an individual when the predisposition for that activity has been activated (Hidi & Renninger, 2006). By contrast, situational interest emerges from specific conditions and task features in the environment that focus on the individual. As a result, situational interest generally invokes an immediate interactive reaction to environmental stimuli. Still, this interest form may also increase knowledge and value as these activities are constantly re-engaged over time (Hidi & Harackiewicz, 2000).

In addition to interest, researchers find that self-efficacy shapes educational outcomes. Students who receive negative feedback regarding their academic abilities are inclined to view themselves as less intelligent. In contrast, those who receive positive feedback are more inclined to view themselves as intelligent. Students who lowered their self-efficacy were less prone to take advanced math courses (Crosnoe et al., 2007).

Researchers emphasizing mathematical differences among men and women often cite self-efficacy as a critical source for these differences (Correll, 2001). These differences start modestly in the early educational years (Hyde, Fennema, & Lamon, 1990; Hyde, Fennema, Ryan, et al., 1990) and become more apparent in high school and beyond (Hyde, Fennema, Ryan, et al., 1990; Leahey & Guo, 2001). When boys and girls perceive their math self-efficacy to be congruent to societal expectations, their performance reflects that perception. However, when individuals perceive no differences, boys and girls perform at equal rates on evaluative measures (Correll, 2001).

This positive (negative) effect of self-efficacy inflates (deflates) students' opinions and capabilities of themselves and, in turn, increases (decreases) their interest and later options, such as postsecondary and occupational attainment (Leahey & Guo, 2001). Individuals with lower ability, whether actual or perceived, have fewer options and position themselves in a more complicated scenario for success, where they are less likely to attend higher education or funnel into two-year colleges. Therefore, it is not surprising that high-ability students are more likely to participate in a four-year university and are more successful in those institutions than are their peers. A similar pattern should hold for choosing a college major. Individuals with high levels of interest and self-efficacy for a particular subject are more likely to select a major congruent with those interests than those with low levels of interest and self-efficacy.

DATA AND METHODS

This study used U.S. data from the High School Longitudinal Study of 2009 (HSLS:), a nationally representative sample of ninth-grade students surveyed in 2009. Investigators surveyed students again in 2012 (11th grade) and 2016 (3 years after high school). The final sample size is 17,340 respondents, 8,840 of whom attended a four-year college. (Numbers are rounded to the nearest ten due to the National Center for Education Statistics requirements.)

Dependent Variable

The dependent variable is students' field of study among four-year graduates. Specifically, the author used the first major respondents declared during the second follow-up interview (2016), and college majors were measured in two distinct ways. First, the author coded college major as a dummy variable that captures science, technology, engineering, and mathematics (STEM) (coded as 1) from other fields (coded as 0). The STEM majors consist of natural science (e.g., biological sciences), technical–quantitative (e.g., engineering, physical sciences), and health.

While this classification represents a vital college major distinction, it masks important differences within each classification. Second, the author coded college major into eight fields of study: natural science, technical–quantitative, health, business, social science, education, service majors, and letters (reference category). This approach allows for nuanced differences in the influence of precollege factors on college major intention. However, these analyses were restricted to those who attended a four-year college or university. Examining only those who attend a four-year institution may bias the influence of precollege factors on college major intention because of differential selection into four-year colleges and universities.

Socio-Demographic Variables

Students who are academically prepared systematically differ from their less-prepared counterparts. Without considering these differences, estimates of academic preparation on college major selection may not be due to academic preparation but rather to other factors related to both preparation and college major selection. As shown in Table 1 (Appendix), this study includes control indicators to help mitigate bias in estimates of academic preparation. These include race, sex, family background, occupational expectations, and academic achievement. The categories for race were White (reference category), Black, Latino, Asian, Native Americans, and multiracial. The author coded sex as a dummy variable (female = 1, male = 0). Parental education captures the highest education of either parent as four categories: no college (reference group), some college, bachelor's degree, and an advanced degree. Family income is the total family income where each unit is 10 thousand dollars.

Aspirations, Effort, Self-Efficacy, and Interest

The independent variables represent students' aspirations, effort, self-efficacy, and interest in math and science. Whether a student expects an occupation in STEM at age 30 denotes future occupational expectations. The HSLS:09 data set contains several items representing students' effort, self-efficacy, and interest. The author used factor analysis to determine the number of factors that underlie these items. Before determining the number of factors, the author performed a Kaiser–Meyer–Olkin (KMO) test to determine the sampling adequacy for factor analysis. The KMO score was 0.87, indicating adequate sampling (Glen, 2016; Stata Press, 2021b). The author selected a three-factor solution based on eigenvalues, scree plot, and conceptual sense. Science self-efficacy is a five-item scale ($\alpha = 0.91$) that includes how confident a student: can understand the science textbook, can master skills taught in the science class, does an excellent job on science tests, does an excellent job on science assignments, and enjoys the science class. The author used an equivalent scale for math self-efficacy ($\alpha = 0.90$). Effort is a six-item scale ($\alpha = 0.77$) based on whether a student: paid attention to the math(science) teacher, turned in math(science) assignments on time, and did as little work as possible in the math(science) class (reverse coded). These analyses reveal that self-efficacy is specific to the course subject, whereas effort tends to be more general. All three scales are standardized with a mean of 0 and a standard deviation of 1.

High School Grades and Coursework

The HSLS: 09 provides students' high school transcripts, which contain information about their grade point average (GPA) and coursework. Cumulative GPA is measured on a four-point scale where higher numbers denote higher grades. A series of dummy variables capture the rigor of high school courses across five subjects: English, language, math, science, and technology. Students were considered to have completed a rigorous curriculum for each subject if they took at least 50% of their classes in honors, advanced, college preparatory, Advanced Placement, or International Baccalaureate (yes = 1, otherwise = 0).

College Characteristics

The author included an indicator of the college's academic orientation. U.S. institutions, such as the Massachusetts Institute of Technology, are oriented towards STEM majors, whereas other colleges (e.g., bible colleges and art colleges) are oriented in fields outside of STEM. A college's academic orientation is a selection measure because enrolling in a specialized school reflects a student's initial commitment to an area of study. Therefore, the type of major students pursue is partially dependent on the institution they attend. The author included a measure of the percent of STEM degrees conferred at an institution in 2012. This measure was created from the Integrated Postsecondary Education Data System (2021), and each unit represents a ten percentage-point change.

Missing Data

The author used multiple imputation techniques to handle missing information. This approach creates $M > 1$ sets of imputed values by introducing random variation to the imputation procedure, creating M slightly different versions of the complete data (Collins et al., 2001). Compared to other treatments of handling missing information, this approach tends to produce larger standard errors due to the introduction of between-imputations variability to its calculation of standard errors. There were a total of 50 replications of the data created.

Method of Analysis

The author used probit regression with sample selection when considering whether students chose a STEM major. Not all individuals attend four-year institutions, and examining only those who participated in four-year institutions may bias the influence of academic accomplishments (and family background) on college major selection. Students who do not attend a four-year college are likely to differ systematically from students who attend four-year institutions. To address the issue of sample selection, the probit selection model jointly estimates the likelihood of attending a four-year college or university and college major selection. Formally, the probit selection model (Stata Press, 2021a; Van de Ven & Van Praag, 1981) assumes that an underlying regression relationship exists:

$$Y_{1i}^{*} = \beta X_i + \varepsilon_i \tag{1}$$

in such a way that individuals only observe the binary outcome if $Y_{1i}^{p} = (Y_{1i}^{*} > 0)$. For the ***i***th individual, let Y_{1i}^{*} represent a latent variable, β represents a vector of coefficients corresponding to a vector of independent variables X_i, and ε_i is the error term. However, the dependent variable is only observed for those who attended a four-year college:

$$Y_{2i}^{S} = aZ_i + v_i > 0. \tag{2}$$

Let Y_{2i}^{S} represent a latent continuous variable that represents the propensity for a student to enroll in a four-year college. Moreover, a represents parameters corresponding to the explanatory variables Z_i, and v_i is the error term. The estimated ρ (rho)—the correlation between ε and v—indicates whether selection on unobservables is an issue ($\rho \neq 0$). If ρ is not statistically significant ($\rho = 0$), then selection on unobservables is assumed to influence minimally equation 1, and therefore the joint estimation of equations 1 and 2 are not required. To help with identification, the author included eleventh-grade expectations of graduating from a four-year college (yes =1, no = 0), the distance to the nearest college that offers a bachelor's degree (one unit = 10 miles), and the distance to the nearest college that offers an advanced degree (one unit = 10 miles) into the selection equation.

The author used multinomial logistic regression for analyses of the dependent variable with multiple categories. Interpretation of both probit selection and multinomial logistic models can be cumbersome. Therefore, the author reported average marginal effects, which converts probit or log-odds scores to the probability scale (Bartus, 2005). The study also weighed the observations and accounted for the sampling design to adjust for oversampling, attrition, and non-response.

RESULTS

Majoring in STEM Fields

Table 2 (Appendix) shows results from the outcome equation of the probit selection model. Model 1 contains the estimated coefficients of socio-demographic characteristics. The estimated rho is –0.48 and is statistically significant, suggesting that models that do not capture self-selection lead to biased estimations. Although not the paper's primary focus, researchers may be interested in the importance of socio-demographic factors on college major choice. The results from Model 1 show that racially minoritized students are at least as likely to select a STEM major as White students. The exception is that Asian American students are 12.2 percentage points higher in their probability of selecting a STEM major than White students.

Model 1 further shows that female students are 3.8 percentage points lower in their probability of selecting a STEM major than male students. Surprisingly, affluent students—as measured by parental education and family income—are less likely to select a STEM major than students whose parents did not go to college. For example, students with at least one parent who attained a baccalaureate are 5.4 percentage points less likely to select a STEM major than students with parents who did not go to college. These results are consistent with previous studies that show students from affluent backgrounds may forego their initial earnings after graduating from college. Instead, they tend to gravitate towards majors (e.g., liberal arts) that lead into graduate school (Davies & Guppy, 1997; Goyette & Mullen, 2006). Not surprisingly, as the percentage of bachelor's degrees conferred in STEM increases at an institution, students' likelihood of choosing a STEM major also increases.

Model 2 includes occupational expectation, self-efficacy, and effort in math and science. Students who expect to work at a STEM occupation by age 30 are 35.8 percentage points higher in their probability to major in STEM than similar students who do not expect this type of occupation. Perhaps not surprisingly, both science and math self-efficacy increases the likelihood of majoring in STEM. For example, a standard deviation increase in science self-efficacy increases the probability of majoring in STEM by 5.3 percentage points. What is surprising is the inverse relation between effort in math and science, and selecting a STEM major. In other words, as students increase their effort in math and science, they are less likely to major in STEM.

After controlling for STEM aspirations, self-efficacy, and effort in math and science, the gender discrepancy in majoring in STEM fields is reduced by 27 percent, from –0.038 (Model 1) to –0.028 (Model 3), and it is marginally significant ($\boldsymbol{p} < 0.10$). Moreover, aspirations, self-efficacy, and effort account for the parental-education difference in STEM majors selection. For example, the difference in STEM majors between students whose parents did not attend college and whose parents' highest education is a bachelor's degree is reduced by 81.5%, from 0.05 in Model 1 to 0.01 in Model 2, and is no longer statistically significant. The effect of family income on selecting a STEM major is reduced by 37.4%, from Model 1 to Model 2; however, this effect remains statistically significant. Interestingly, aspirations, self-efficacy, or effort does little to explain the Asian American advantage in STEM major selection.

Model 3 includes students' high school grades and coursework. A one-unit change in GPA is associated with a 10.5 percentage-point increase in the probability of selecting a STEM major. Moreover, rigorous math and science courses increase students' probability of selecting a STEM major by 6.1 and 3.3 percentage points, respectively. Surprisingly, students taking rigorous English, language, or even technology courses do not influence their probability of selecting a STEM major. Including grades and coursework reduces the relation between effort and selecting a STEM major by 47.3%, from –0.032 in Model 2 to –0.017 in Model 3, and is no longer statistically significant.

Furthermore, grades and coursework reduce the influence of self-efficacy in math and science on selecting a college major by 18.1% and 29.1%, respectively. Despite these reductions in their effects, science and math self-efficacies remain statistically significant. Even in Model 3, Asian American students continue to select STEM majors at higher rates than White students, although the effect was reduced by 24.5%. This result implies that Asian Americans majoring in STEM fields are not entirely due to differences in their academic accomplishments.

Students' Choices across Fields of Study

Tables 3 (Appendix) presents the final results from the multinomial regression analysis. The purpose of this analysis is to examine whether the binary outcome used in the probit selection model masks nuanced differences within the STEM and non-STEM categories. However, the downside is that the multinomial regression analysis only examines those who attended a four-year institution. It does not account for sample selection as with the probit selection models in the previous analyses.

All things being equal, racially minoritized students are at least as likely to select STEM majors as White students. The exception is that Asian American students are 5 percentage points more likely to select technical–quantitative majors than White students. Moreover, Asian

American and multiracial students are 6.7 and 2.8 percentage points, respectively, more likely to select a natural science major than White students. Black, Latino, and multiracial students are 8.3, 3.8, and 3.4 percentage points, respectively, more likely to select a social science major than White students. There are few differences between racially minoritized students and White students in selecting business, service, and letters. The exception is that Black and Asian American students are both 3.3 percentage points less likely to select a major in letters than White students. Moreover, racially minoritized students are less likely to select a major in education than White students; although the difference between Native American and White students are marginally significant ($p < 0.10$).

Female students tend to display different major patterns than male students, even after accounting for family background, aspirations, self-efficacy, effort, and academic accomplishments. In general, female students tend to select majors in health, social science, and education at higher rates (10.5, 5.2, and 5.8 percentage points, respectively), but select technical–quantitative and business (14.9 and 7.3 percentage points, respectively) at lower rates than male students. However, female students are as likely as male students to select natural science, service, and letters.

Results show some evidence that college major decisions fall along social class divisions. Students with a parent who attained at least a baccalaureate are more likely to select a technical–quantitative major but less likely to select a service major. Moreover, students with a parent who attained an advanced degree are 2.1 percentage points less likely to major in health than those whose parents did not attend college. There is little difference between parental education and majoring in natural science, business, social science, education, service, and letters. Affluent students—based on family income—are more likely to select business and social science majors. Moreover, they are less likely to select an education major. Interestingly, selecting health and technical–quantitative majors are reserved for less-affluent students.

Students who expect to work in a STEM-related occupation by age 30 are associated with college major decisions. In other words, this occupational expectation increases the probability of majoring in natural science, technical–quantitative, and health while decreasing the probability of majoring in business, social science, education, and letters. Not surprisingly, the expectations of working in a STEM-related occupation by age 30 are incredibly influential for students majoring in health. The probability of selecting natural science and technical–quantitative majors are 9.5 and 6.8 percentage points, respectively, for students who expect to work in a STEM-related occupation by age 30. By contrast, this occupational expectation increases the probability of selecting a health major—majors more closely aligned with professional occupations—by 15 percentage points.

As expected, a standard-deviation increase in science and math self-efficacy increases the probability of selecting a technical–quantitative major by 2.8 and 6.3 percentage points, respectively. However, math self-efficacy does not affect students' probability in selecting natural science, but science self-efficacy does increase their probability of selecting these majors. Conversely, science self-efficacy does little to predict selection into health majors, whereas a standard-deviation increase in math self-efficacy decreases the probability of majoring in a health major by 1.6 percentage points. Effort in math and science negatively influences those who select a technical–quantitative major, but not for the other STEM majors.

High school GPA is positively associated with students selecting a major in natural sciences and technical–quantitative. Students with good high school GPAs are less likely to select a major in business and service. There is no association between high school GPA and college major decisions in education, health, social science, education, or letters.

High school course-taking patterns contribute to students' undergraduate major choices. For example, students with rigorous high school coursework in math, science, or technology are more likely to select a technical–quantitative major than those with less rigorous course-taking patterns. This relation between STEM courses in high school and a STEM undergraduate major is even more pronounced for students who took a rigorous technology course load in high school. Moreover, students who took a rigorous English or language course load in high school are less likely to select a technical–quantitative major. Interestingly, a rigorous science course load in high school decreases students' probability of majoring in health by 1.8 percentage points. Furthermore, a rigorous English course load in high school increases students' probability of majoring in social science and letters by 3.6 and 3 percentage points, respectively. A rigorous English course load decreases the likelihood of selecting some professional majors such as business and service.

DISCUSSION AND CONCLUSION

This paper looked at how the transference of expectations, self-efficacy, effort, and preparation in high school sorts students into different fields of study. The results exhibited consistencies, extensions, and discrepancies with past research. Consistent with previous studies, this analysis generally found little difference among Black and Latino students' college major preferences compared to White students, which supported work by Simpson (2001). Moreover, the author found that Asian American students were inclined to major in science-oriented and quantitative fields. The author also found evidence to support Wilson and Boldizar's (1990) claim that women may seek to improve their social standing through means outside non-math fields such as health. The results further indicate that the association between gender and STEM fields found in the probit selection model is due to offsetting forces of the overrepresentation of women in health-related fields and the underrepresentation of women in technical fields—and to some extent—the overrepresentation of women in education and the underrepresentation of women in business.

At first, high family income lowered students' likelihood of majoring in non-STEM fields. However, this association was due to students' differences in high school coursework across the family-income distribution. Upon closer inspection, students from families with high family incomes tend to select business and social science majors. The lack of association between parental association and selection of a STEM major may be due to the offsetting effects where students with a parent who attained at least a baccalaureate are more likely to select a technical–quantitative major but less likely to select a health major. The author found little evidence that students with a parent who attained an advanced degree are more likely to major in letters. This result is inconsistent with Goyette and Mullen's (2006) finding that students from affluent backgrounds may forego their initial earnings after college to pursue a graduate degree.

Students' occupational expectations, self-efficacy, and effort in STEM during high school are important indicators for their behavior in college. On the one hand, students' expectations of working at a STEM occupation by age 30 are especially consequential for selecting a health major. On the other hand, students' self-efficacy in math and science is more important for selecting a technical–quantitative major. Interestingly, strenuous effort in math or science reduces the likelihood of selecting a technical–quantitative major. This result suggests that students' effort reflects less grit and more "survival" in these subjects.

This paper supported past studies that claim the importance of precollege science indicators for students choosing science-related majors. However, this analysis differed from previous studies that examined the relation between academic preparation and college major selection. The author emphasized academic rigor by using subject-specific pipeline measures instead of a single composite measure of academic intensity (Adelman, 2006) or measures representing the number of subject-specific high school credits a student took (Simpson, 2001; Wang, 2013). In particular, students taking rigorous math, science, or technology course load in high school was especially consequential in majoring in a technical–quantitative major. Similarly, English coursework was significant for students' decision to major in letters. Despite what many presuppose, precollege math coursework was not a substantial predictor for students majoring in natural science or health.

This paper advocates the importance of precollege preparation and academic interest in influencing students' college access and experiences. Additionally, this paper provides evidence (or lack thereof) and offers limitations of the extent to which previous concepts and findings can apply. However, more research is needed regarding college majors and students' undergraduate tendencies. More research is also needed to explain Asian American students' tendency to major in science-related fields. Xie and Goyette (2003) lay a framework where researchers can pursue this advantage. Research is further needed to explain the underrepresentation of women in quantitative fields as well as their overrepresentation in health.

REFERENCES

Adelman, C. (2006). ***The toolbox revisited: Paths to degree completion from high school through college***. U.S. Department of Education. https://www2.ed.gov/rschstat/research/pubs/toolboxrevisit/index.html

An, B. P. (2013a). The impact of dual enrollment on college degree attainment: Do low-SES students benefit? ***Educational Evaluation and Policy Analysis***, ***35***(1), 57–75. https://doi.org/10.3102/0162373712461933

An, B. P. (2013b). The influence of dual enrollment on academic performance and college readiness: Differences by socioeconomic status. ***Research in Higher Education***, ***54***(4), 407–432. https://doi.org/10.1007/s11162-012-9278-z

Bartus, T. (2005). Estimation of marginal effects using margeff. ***Stata Journal***, ***5***(3), 309–329.

Behrman, J. R., Rosenzweig, M. R., & Taubman, P. (1996). College choice and wages: Estimates using data on female twins. ***Review of Economics and Statistics***, ***78***(4), 672–685. https://doi.org/10.2307/2109954

Burton, M. B. (1989). The effect of prior calculus experience on "introductory" college calculus. ***American Mathematical Monthly***, ***96***(4), 350–354. https://doi.org/10.2307/2324093

Charles, M., & Bradley, K. (2002). Equal but separate? A cross-national study of sex segregation in higher education. ***American Sociological Review***, ***67***(4), 573–599. https://doi.org/10.2307/3088946

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. ***Psychological Methods***, ***6***(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. ***American Journal of Sociology***, ***106***(6), 1691–1730.

Crosnoe, R., Riegle-Crumb, C., & Muller, C. (2007). Gender, self-perception, and academic problems in high school. ***Social Problems***, ***54***(1), 118–138. https://doi.org/10.1525/sp.2007.54.1.118

Davies, S., & Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. ***Social Forces***, ***75***(4), 1417–1438. http://links.jstor.org/sici?sici=0037-7732%28199706%2975%3A4%3C1417%3AFOSCSA%3E2.0.CO%3B2-W

Duta, A., An, B., & Iannelli, C. (2018). Social origins, academic strength of school curriculum and access to selective higher education institutions: Evidence from Scotland and the USA. ***Higher Education***, ***75***(5), 769–784. https://doi.org/10.1007/s10734-017-0166-5

Engberg, M., & Wolniak, G. (2013). College student pathways to the STEM disciplines. ***Teachers College Record***, ***115***(1), 1–27. https://doi.org/10.1177/016146811311500102

Ferrini-Mundy, J., & Gaudard, M. (1992). Secondary school calculus: Preparation or pitfall in the study of college calculus? ***Journal for Research in Mathematics Education***, ***23***(1), 56–71. https://doi.org/10.5951/jresematheduc.23.1.0056

Gerber, T. P., & Cheung, S. Y. (2008). Horizontal stratification in postsecondary education: Forms, explanations, and implications. ***Annual Review of Sociology***, ***34***(1), 299–318. https://doi.org/10.1146/annurev.soc.34.040507.134604

Glen, S. (2016). ***Kaiser–Meyer–Olkin (KMO) test for sampling adequacy***. StatisticsHowTo.com. Retrieved November 18 from https://www.statisticshowto.com/kaiser-meyer-olkin/

Goyette, K. A., & Mullen, A. L. (2006). Who studies the arts and sciences? Social background and the choice and consequences of undergraduate field of study. ***Journal of Higher Education***, ***77***(3), 497–538. https://doi.org/10.1353/jhe.2006.0020

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. ***Journal of Educational Psychology***, ***92***(2), 316–330. https://doi.org/10.1037/0022-0663.92.2.316

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. ***Journal of Educational Psychology***, ***94***(3), 562–575. https://doi.org/10.1037/0022-0663.94.3.562

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. ***Review of Educational Research***, ***70***(2), 151–179. https://doi.org/10.3102/00346543070002151

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. ***Educational Psychologist***, ***41***(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance : A meta-analysis. ***Psychological Bulletin***, ***107***(2), 139–155. https://doi.org/10.1037/0033-2909.107.2.139

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. ***Psychology of Women Quarterly***, ***14***(3), 299–324. https://doi.org/10.1111/j.1471-6402.1990.tb00022.x

Integrated Postsecondary Education Data System. (2021). ***Use the data***. National Center for Education Statistics. Retrieved October 18 from https://nces.ed.gov/ipeds/use-the-data

Karen, D., & Dougherty, K. J. (2005). Necessary but not sufficient: Higher education strategy of social mobility. In G. Orfield, P. Marin, & C. L. Horn (Eds.), ***Higher education and the color line: College access, racial equity, and social change*** (pp. 33–57). Harvard Education Press.

Kerckhoff, A. C. (2001). Education and social stratification processes in comparative perspective. ***Sociology of Education***, ***74***(Extra Issue), 3–18. https://doi.org/10.2307/2673250

Kim, C., Tamborini, C. R., & Sakamoto, A. (2015). Field of study in college and lifetime earnings in the United States. ***Sociology of Education***, ***88***(4), 320–339. https://doi.org/10.1177/0038040715602132

Kirst, M. W., & Bracco, K. R. (2004). Bridging the great divide: How the K–12 and postsecondary split hurts students, and what can be done about it. In M. W. Kirst & A. Venezia (Eds.), ***From high school to college: Improving opportunities for success in postsecondary education*** (pp. 1–30). Jossey-Bass.

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. ***Journal for Research in Mathematics Education***, ***32***(5), 448–470. https://doi.org/10.2307/749801

Lareau, A. (2011). ***Unequal childhoods: Class, race, and family life (2nd Edition)***. University of California Press.

Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. ***Social Forces***, ***80***(2), 713–732. https://doi.org/10.7227/RIE.0015

Li, H., Liu, P. W., & Zhang, J. (2012). Estimating returns to education using twins in urban China. ***Journal of Development Economics***, ***97***(2), 494–504. https://doi.org/10.1016/j.jdeveco.2011.05.009

Lucas, S. R. (1999). ***Tracking Inequality: Stratification and mobility in American high schools***. Teachers College Press.

Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. ***American Journal of Sociology***, ***106***(6), 1642–1690. https://doi.org/10.1086/321300

Lucas, S. R. (2017). An archaeology of effectively maintained inequality theory. ***American Behavioral Scientist***, ***61***(1), 8–29. https://doi.org/10.1177/0002764216682989

Parsad, B., & Lewis, L. (2003). ***Remedial education at degree-granting postsecondary institutions in fall 2000***. National Center for Education Statistics

Posselt, J. R., Jaquette, O., Bielby, R., & Bastedo, M. N. (2012). Access without equity: Longitudinal analyses of institutional stratification by race and ethnicity, 1972–2004. ***American Educational Research Journal***, ***49***(6), 1074–1111. https://doi.org/10.3102/0002831212439456

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into stem college majors over time. ***American Educational Research Journal***, ***49***(6), 1048–1073. https://doi.org/10.3102/0002831211435229

Russell, M. L., & Atwater, M. M. (2005). Traveling the road to success: A discourse on persistence throughout the science pipeline with African American students at a predominantly White institution. ***Journal of Research in Science Teaching***, ***42***(6), 691–715. https://doi.org/10.1002/tea.20068

Schiller, K. S., & Hunt, D. J. (2011). Secondary mathematics course trajectories: Understanding accumulated disadvantages in mathematics in grades 9–12. ***Journal of School Leadership***, ***21***(1), 87–118. https://doi.org/10.1177/105268461102100105

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***(1), 63–100. https://doi.org/10.2307/2649134

Stata Press. (2021a). ***Stata base reference manual (release 17)***. StataCorp.

Stata Press. (2021b). ***Stata multivariate statistics reference manual (release 17)***. StataCorp.

Stearns, E., Jha, N., & Potochnick, S. (2013). Race, secondary school course of study, and college type. ***Social Science Research***, ***42***(3), 789–803. https://doi.org/10.1016/j.ssresearch.2013.01.007

Stevenson, D. L., Schiller, K. S., & Schneider, B. (1994). Sequences of opportunities for learning. ***Sociology of Education***, ***67***(3), 184–198. https://doi.org/10.2307/2112790

Tamborini, C. R., Kim, C., & Sakamoto, A. (2015). Education and lifetime earnings in the United States. ***Demography***, ***52***(4), 1383–1407. https://doi.org/10.1007/s13524-015-0407-0

Tyson, K. (2011). ***Integration interrupted: Tracking, black students, and acting white after Brown***. Oxford University Press.

Useem, E. L. (1992). Middle schools and math groups: Parents' involvement in children's placement. ***Sociology of Education***, ***65***(4), 263–279. http://links.jstor.org/sici?sici=0038-0407%28199210%2965%3A4%3C263%3AMSAMGP%3E2.0.CO%3B2-V

Van de Ven, W. P. M. M., & Van Praag, B. M. S. (1981). The demand for deductibles in private health insurance: A probit model with sample selection. ***Journal of Econometrics***, ***17***(2), 229–252. https://doi.org/10.1016/0304-4076(81)90028-2

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. ***American Educational Research Journal***, ***50***(5), 1081–1121. https://doi.org/10.3102/0002831213488622

Webber, D. A. (2016). Are college costs worth it? How ability, major, and debt affect the returns to schooling. ***Economics of Education Review***, ***53***, 296–310. https://doi.org/10.1016/j.econedurev.2016.04.007

Wilson, K. L., & Boldizar, J. P. (1990). Gender segregation in higher education: Effects of aspirations, mathematics achievement, and income. ***Sociology of Education***, ***63***(1), 62–74. https://doi.org/10.2307/2112897

Xie, Y., & Goyette, K. (2003). Social mobility and the educational choices of Asian Americans. ***Social Science Research***, ***32***(3), 467–498. https://doi.org/10.1016/S0049-089X(03)00018-8

APPENDIX

Table 1. Description of Variables

Variables	Description	Mean	S.D.
Dependent variables			
STEM Majors	The first major or field of study declared or decided upon is in a STEM field (yes = 1)	0.40	0.49
Field of Study	The first major or field of study declared		
	Natural science	0.09	
	Technical–quantitative	0.16	
	Health	0.16	
	Business	0.16	
	Social science	0.13	
	Education	0.06	
	Service majors	0.14	
	Letters (omitted)	0.10	
Control variables			
Background			
Black	Student is Black	0.10	0.30
Latino	Student is Latino	0.16	0.36
Asian	Student is Asian	0.09	0.29
Native American	Student is Native American	0.01	0.09
Multiracial	Student is Multiracial (White is omitted)	0.09	0.28
Female	Student is female	0.51	0.50
Some college	At least one parent attended some college	0.21	0.41
Bachelor's degree	At least one parent earned a bachelor's degree	0.25	0.43
Advanced degree	At least one parent earned an advacned degree (no college is omitted)	0.20	0.40
Family income	Average family income from all sources (10 thousand)	8.54	6.04
Degrees conferred at college			
STEM degrees conferred	BA degrees conferred in STEM (1 = 10% change) (2012)	2.01	1.26
Occupational expectations, self-efficacy, and effort			
Future STEM occupation	Expects a STEM occupation at age 30 (yes = 1)	0.35	0.48
Science self-efficacy	Five items ($\alpha = 0.91$) (see text for items)	-0.01	0.82

Math self-efficacy	Five items (α = 0.90) (see text for items)	-0.01	0.81
Effort in math or science	Six items (α = 0.77) (see text for items)	-0.04	0.68
Academic achievement and coursework			
Grade point average (GPA)	Cumulative GPA. Four-point scale (4 = "A")	2.83	0.70
Rigorous coursework (English)	50% or more English courses are honors (1 = Yes)	0.37	0.48
Rigorous coursework (language)	50% or more language courses are honors (1 = Yes)	0.16	0.37
Rigorous coursework (math)	50% or more math courses are honors (1 = Yes)	0.32	0.47
Rigorous coursework (science)	50% or more science courses are honors (1 = Yes)	0.32	0.47
Rigorous coursework (tech.)	50% or more technology courses are honors (1 = Yes)	0.02	0.12
Instruments			
Expectations to earn BA	11th-grade expectation of earning a BA or higher (1 = Yes)	0.67	0.46
Distance to nearest college (BA)	Distance to nearest BA college using zipcode of high school (miles)	315.64	406.33
Distance to nearest college (advanced)	Distance to nearest MA/Doctoral College using zipcode of high school (miles)	179.46	238.13

Note: HSLS:09. Sample size is 16,840, 8,340 of whom attended a four-year college. Numbers are rounded to the nearest 10. Means and standard deviations are unweighted.

Table 2. Probit Selection Models Estimating the Likelihood of Majoring in STEM Fields

	Model 1	Model 2	Model 3
	ME	ME	ME
Background			
Black	.003	–.029	.015
Latino	.027	.005	.007
Asian	.122***	.116***	.087**
Native American	.119	.022	.034
Multiracial	.032	–.003	.003
Female	–.038*	–.028†	–.027*
Some college	–.049†	–.036	–.017
Bachelor's degree	–.054**	–.01	.026
Advanced degree	–.042†	–.002	.035†
Family income	–.006***	–.004*	–.001
Occupational expectations, self-efficacy, and effort			
Future STEM occupation		.358***	.328***
Science self-efficacy		.053***	.043***
Math self-efficacy		.066***	.046***
Effort in math or science		–.032†	–.017
Academic achievement and coursework			
Grade point average (GPA)			.105***
Rigorous coursework (English)			–.012
Rigorous coursework (language)			–.021
Rigorous coursework (math)			.061***
Rigorous coursework (science)			.033*
Rigorous coursework (tech.)			.087
Degrees conferred at college			
STEM degrees conferred	.073***	.067***	.052***
rho	–.476***	–.172**	.331***

Note: HSLS:09. Sample size is 16,840, 8,340 of whom attended a four-year college. Numbers are rounded to the nearest 10. ME = marginal effects. Regressions are weighted, and adjusted for sampling design. † p < .10, * p < .05, ** p < .01, *** p < .001 (two-tailed).

Table 3. Multinominal Logistic Regression of College Major Selection (Marginal Effects)

Variables	Natural	Technical	Health	Business	Social	Education	Service	Letters
Background								
Black	.009	–.016	–.004	.001	.083***	–.04***	.000	–.033**
Latino	.001	.011	–.013	–.004	.038**	–.036***	.016	–.014
Asian	.067***	.050***	–.018†	.004	–.002	–.052***	–.016	–.033***
Native American	.020	.031	–.015	–.074	.005	–.049†	.092	–.011
Multiracial	.028*	–.015	–.004	–.008	.034*	–.04***	.010	–.005
Female	.008	–.149***	.105***	–.073***	.052***	.058***	–.01	.008
Some college	–.012	.008	.001	.000	.002	–.002	.021†	–.019†
Bachelor's degree	.005	.023†	–.014	.012	.004	–.009	–.012	–.009
Advanced degree	.010	.035**	–.021*	–.018	.011	–.014	–.014	.010
Family income	.000	–.001*	–.001*	.003***	.002***	–.002***	–.001†	.000
Occupational expectations, self-efficacy, and effort								
Future STEM occupation	.095***	.068***	.150***	–.114***	–.054***	–.041***	–.014†	–.09***
Science self-efficacy	.026***	.028***	–.006	–.036***	–.007	–.003	–.006	.004
Math self-efficacy	.010†	.063***	–.016***	.001	–.028***	.000	–.008†	–.021***
Effort in math or science	.013†	–.043***	.016*	.021**	–.007	.012*	.000	–.013*
Academic achievement and coursework								
Grade point average (GPA)	.025**	.036***	–.013†	–.027**	.011	.003	–.038***	.003
Rigorous coursework (English)	.000	–.031***	.006	–.022*	.036***	–.002	–.017*	.030***
Rigorous coursework (language)	.004	–.019*	–.01	.015	.017	–.006	–.016†	.014
Rigorous coursework (math)	.002	.050***	–.007	.011	–.035***	–.001	.000	–.02*
Rigorous coursework (science)	.016†	.034***	–.018*	–.009	.003	–.016*	–.012	.002
Rigorous coursework (tech.)	–.08***	.124***	–.033	.009	.007	–.004	.021	–.043*
Degrees conferred at college								
STEM degrees conferred	.011**	.035***	.019***	–.013**	–.009†	–.014***	–.006	–.024***

Note: HSLS:09. Sample size is 8,190. Numbers are rounded to the nearest 10. Regressions are adjusted for sampling design. † p < .10, * p < .05, ** p < .01, *** p < .001 (two-tailed).