STEM MAJOR CHOICE: HIGH SCHOOL AND COLLEGIATE FACTORS

by

Que Nguyet Tran

A dissertation submitted in partial fulfillment

of the requirements for the degree

of

Doctor of Philosophy

in

Adult and Higher Education

MONTANA STATE UNIVERSITY

Bozeman, Montana

May 2022

DEDICATION

To my parents Tran Cau and Vy Tuong Nga,
sister Tran Thanh Hai,
nieces Dang Tran Vi Linh, Dang Tran Linh Anh, Tran Hai Anh,
and son Lam Gia Bach!

ACKNOWLEDGEMENTS

"It takes a village", I owe an immeasurable debt of many people who have supported me through the dissertation. I would like to express my deepest gratitude to my advisor, Dr. Tricia Seifert, I am eternally grateful for your guidance, mentorship, patience, care, and support! Words cannot capture the magnitude of appreciation I have for my committee members: Drs. Carrie B. Myers, Fenqjen Luo, and Bryce Hughes. Your dedication to student development, and most importantly, being epitomized what it means to be a scholar, and educator that I'll mirror in my next career chapter. A special thanks for Deans Harmon and Ogilvie for your trust in my success. I am indebtedly thankful for EHHD faculty and other MSU faculty who help me develop educational knowledge, statistical, and English academic writing skills, your support have never gone unnoticed through my scholarship journey. Drs. Alexandra Aylward and Mary Kate Blake, thank you for provoking me to critically engage within inquiries, and graduate-level instructions. I also would like to express my appreciation for OIP team who have always made me feel home. Thank you to my dear friend Catherine Johnsons who expanded my horizon geographically and scholarly. A sweet memory of James Sletten who would have celebrated with me. Thank you, my friends, for your warmth and kindheartedness, Robin Clausen, Mr. and Mrs. Bret and Becky McCotter, Mr. and Mrs. John and Deb Sherer. I owe a huge thank to my dearest sister, friend, mentor, Dr. Catherine Yap Co for your endless love, I am more thankful than you'll ever know. Finally, thank you for the Hubert H. Humphrey Fellowship to bring me the opportunity to become a researcher, and friendships near and far.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

Table Page

LIST OF TABLES CONTINUED

LIST OF FIGURES

ABSTRACT

A huge present and future workforce demand exists in Science, Technology, Engineering, and Mathematics (STEM) fields. Bolstered by a number of US policies and research that associates STEM majors with pursuing STEM careers, higher education institutions have aimed to support students to major in STEM fields in an effort to meet the needs of the STEM workforce. Despite these postsecondary efforts, the challenge begins in earlier levels of schooling with a shortage of licensed and highly qualified science and math teachers nationwide. Although many studies have examined math and science expectancy values and self-efficacy among high school students to predict their intention to major in STEM major choice, few have investigated both high school and college level variables to understand student STEM major choice declared in their third college year. Thus, this study fills the gap using the most recent STEM-focused national representative survey data – High School Longitudinal Study 2009 (HSLS:09). Three research questions are: (i) To what extent do high school math and science motivation and self-efficacy, collegiate factors, and personal circumstances promote or hinder students' STEM major choice, controlling for student background characteristics?; (ii) To what extent do collegiate factors and personal circumstances predict the probability of STEM major choice, controlling for student background characteristics? (iii) What factors predict college STEM GPA? This study employs theories of Situative expectancy value theory and Social cognitive career theory to develop a conceptual framework. Logistic regression was used to analyze the first two questions, and linear regression used for the third question.

The first research question found gender, math attainment value, science attainment value, college STEM credits earned, and STEM GPA are predictive of the probability of STEM major choice. In the second research question, among college-period variables, gender, college STEM credits earned, and STEM GPA are predictors of STEM major choice. The third research question found race, social economic status, faculty research participation, career services on campus used, work schedule and academic performance interference, and disability are predictors of the average STEM GPA. Implications for theory, research, and practice are discussed.

LIST OF ABBREVIATIONS

BLS	Bureau of Labor Statistics
DOE	Department of Education
GPA	Graduate Point Average
LGBQ	Lesbian, Gay, Bisexual, Queer
NCES	National Center for Education Statistics
NSF	National Science Foundation
SES	Social Economics Status
STEM	Science, Technology, Engineering, Math

CHAPTER ONE

INTRODUCTION

Shortage of STEM workforce

Increasing the number of students who major in Science, Technology, Engineering and Mathematics (STEM) and subsequently earn a bachelor's or higher degree in these fields is a national goal in the US. Several policies have been implemented such as "Educate to Innovate campaign" launched by the former president, Barack Obama, in 2009 (White House, 2009) to expand STEM education and career opportunities for underrepresented groups, including women and minorities (*see Appendix B*). This expansion has been a long-standing goal for many years, and has been, at least in part, accomplished by the growing number of undergraduates pursuing STEM degrees. According to NCES reports (Department of Education, 2019), the growth in STEM bachelors conferred has increased 163% during the ten-year period from 2008-09 (243,031 degrees) to 2017-18 (395,243 degrees). The demands in STEM workforce continue apace despite the increased patterns of undergraduate enrollment in STEM. For example, by 2025, the US is projected to have over two million unfilled STEM jobs (Apriceno et al., 2020) and specifically 3.6 million US computing-related job openings are expected by 2029 (National Center for Women & Information Technology, 2021). Furthermore, the number of STEM jobs is projected to grow by 8% compared to 3.4% for non-STEM jobs in period 2019-2029. Among the STEM jobs, 73% will require a bachelor's degree for entry (Fayer et al., 2017; US Bureau of Labor Statistics, 2020). Filling open STEM jobs remains a challenge.

In order to meet the demand in the STEM workforce, the US must diversify the student population pursuing STEM majors and seeking employment in STEM fields. Many policies have called for more initiatives to increase historically underrepresented students (i.e., women and racialized minorities) in pursuing a STEM college education (Johnson et al., 2020). Women and minorities comprise 70% of the college student population but earn fewer than 45% of STEM degrees awarded (Blackburn, 2017). According to the US Census Bureau (US Census Bureau, 2019) and Catalyst (Catalyst, 2021), one in five Americans is a woman of color and women comprise approximately 51 % of the US population. By 2044, more than half of all Americans are projected to belong to a minority group (Colby & Ortman, 2015).Thus, given this gender and racial distribution in the population, encouraging a broader diversity of college female and underrepresented students to pursue STEM majors would facilitate economic mobility for women and underrepresented students in the 21st century as STEM jobs often have higher earnings than non-STEM students, and potentially to meet the future STEM workforce demand.

Different career aspirations between gender, race/ethnicity

This study assumes that women and students of color seek a STEM degree because they are competent, expected a STEM career to be financially rewarding and personally fulfilling (Patnaik et al., 2020). For example, in information technology industry, according to the National Center for Women in Technology, 57% of professional occupations in the 2020 US workforce are held by women, but only a quarter of the US computing positions in the 2020 US workforce were held by women, and only 18% of the Chief Information Officer (CIO) positions in top 1000 companies were held by women (National Center for Women & Information Technology, 2021). Notably, empirical study showed evidence that firms with female Chief Technology Officers (CTOs) were more innovative than firms with male CTOs, and the effect of female CTOs on corporate innovation was positive for firms with a strong corporate culture (Wu et al., 2021).

A lack of diversity among the creators of innovation would affect economic mobility of women as computing jobs are among the highest paid occupations (Bureau of Labor Statistics, 2021), and by large, affect the country's innovative competitiveness. Although 57% of 2019 bachelor's degree recipients were women, the disproportional gender and race/ethnicity disparities within the computing workforce are wide. Specifically, 37% of Computer Science bachelor's degree receivers in 1985 were women, and this rate dropped to 21% in 2019. Moreover, only 3%, 7%, and 2% were respectively African American, Asian, Hispanic women in the computing workforce in 2020. In 2016, the percentage of Science and Engineering bachelor's degree recipients was respectively 9% and 3.9% among Blacks or African Americans, 13.5% and 10.4% among Hispanic/Latinos, 0.5% and 0.3% among American Indians or Alaska Native, 0.2% and 0.1% among Native Hawaiians or Pacific Islanders, 7.3% and 5.9% among more than one race, 9% and 10.8% among Asians (National Science Foundation, 2019). In engineering, math and statistics, and physical sciences fields, the number of women receiving bachelor's degrees in 2016 were respectively 21% (2.5% increased against 1997), 42.4% (nearly 4% decreased against 1997), and 19.3% (0.1% increased against 1997) (National Science Foundation, 2019). Furthermore, according to the National Science Board, Science and Engineering Indicators 2018, among 6.7 million science and engineering (S&E) workforce, 76% S&E occupations were in computer, mathematical sciences, and engineering with college education. Therefore, it is critical to understand what factors associate with female and underrepresented minority students' decision to major in STEM.

Efforts to support students' exploration pathway in STEM fields need to start early in students' college career. Declaring a STEM major is essential toward a STEM degree that enable women and minority students to bridge to employment realities, cultivate economic mobility, and engage scientific discovery for human advancement (Hoff et al., 2022; Wang & Degol, 2017). In particular, the collegiate supports and personal circumstances as potential factors that may promote or hinder students' retention in the STEM track are worth examining. Yet, students gain accumulated knowledge of math and science in high school in order to prepare for college enrollment and STEM major exploration, however, low persistence rate in STEM major is one of critical reasons that hinder the supply side to meet the STEM workforce demands (Chen, 2013). A large literature has investigated high school academic achievements and STEM college major intention in the first-, or second-year of college (Bolds, 2017; Jiang, Simpkins, et al., 2020; Kurban & Cabrera, 2020; Wang, 2013), or high school academic achievements and STEM graduation rate (Engberg & Wolniak, 2013; Mau, 2016; Wolniak, 2016). Few studies, on the other hand, have examined student college-level STEM academic achievements, the collegiate supports, and personal circumstances in relation to student STEM major choice in the third college year, in connection with high-school math and science achievement-related motivations, math and science self-efficacy. The present study fills the gap.

Push and Pull factors of STEM participation

There are a number of explanations for why students do or do not participate in STEM including psychological, social, cognitive, and career development. These factors are influential at the high school and college levels, Once in college, these factors can then result in a change of major from STEM to another field of study. Within three years of initial enrollment, about 30% of students enrolled in associate's program and bachelor's degree program who already declared a major had changed their major at least once (Department of Education, 2017). Another national report explained reasons why students may have left STEM majors. They included that few high school graduates are academically prepared to enter STEM fields. Few expressed initial interest in entering STEM fields; for some, this was due to the perceived culture of STEM, Once students were off the STEM pathway, it was difficult to get back on. Others expressed concern about earnings from STEM majors compared to other majors (Olson & Rapporteurs, 2012). The finding of the insufficient earnings from STEM majors ten years ago has become irrelevant with the contemporary era.

At high-school level

Students who intended to major in STEM are typically different from all students in precollege preparation as they have better academic preparation at high school and higher average GPA. Past research has shown that math and science knowledge are critical for student participation in STEM postsecondary education (Bolds, 2017; Jiang, Simpkins, et al., 2020; Kurban & Cabrera, 2020; Wang, 2013). Student confidence in math and science during high school are also associated with their intention to a STEM major and contributes to their adolescent career development.

Using High School Longitudinal Study - HSLS:2009 (HSLS:09) dataset, several studies have explained the effects of math and science academic achievements, motivations, and self-efficacy, STEM course-takings at high school (Jiang et al., 2020), AP STEM course exposures (Jewett, 2019), informal STEM learning experiences (Bolds, 2017), and environmental factors such as parental involvement (Kurban & Cabrera, 2020; Mau & Li, 2018), school belonging and math teacher certificates/qualifications (McQuillen, 2016), or math teacher's encouragement to study math (Itauma, 2019) in relation to students' math abilities in high school, and intention to a STEM college major.

Kurban and Cabrera (2020) found student's STEM readiness was associated with GPA in high school STEM courses, credits earned in high school math courses and science courses was affected by math ability measured by student's ability level in math in 9th grade, SES, and parental involvement (Kurban & Cabrera, 2020). Notably, the study utilized structural equation modeling and the model accounted for almost 50% of variance in STEM readiness and 17% of variance in the intention to a STEM major choice. That indicated high school math and science knowledge and skills are critical in student pathway to a STEM college major. Furthermore, Kurban and Cabrera conducted confirmatory factor analysis that showed the reliability of math self-efficacy (omega = .94) and science self-efficacy (omega = .93). Math self-efficacy and science self-efficacy had high impact on math interest and science interest respectively that had direct effects on the intention to a STEM major. Regarding the difference in intention to major in STEM by gender, female students had lower intention than male peers (Kurban & Cabrera, 2020; Mau & Li, 2018). Racially, Asian American students, on average, had higher intention to major in STEM than White students, but there was no difference between underrepresented minority students and White student (Kurban & Cabrera, 2020). Similarly, Mau and Li (2018) applied career aspirations model and found evidence of student background variables (i.e. being male or White – excluding Asian in this study, having a higher SES, greater parental higher expectations) and subject-related variables (i.e. greater math achievement and higher math interest as well as math/science self-efficacy) were significant predictors of the likelihood of STEM career aspirations at age of 30 as students self-reported in comparison with non-STEM students. Moreover, Jewett (2019) reported gender (men more so than women), STEM course credits in high school, AP STEM course exposure, math self-efficacy, science self-efficacy, and math SAT score are predictors of STEM major selection.

At college level

Student interest and confidence in the subject, grades, credits attempted were found to be key factors in pushing students from STEM or pulling them into non-STEM majors. The academic intensity and postsecondary education experiences are more complex than at high school level and these may likely affect student persistence, retention, and graduation in STEM fields. Minaya (2018) used a statewide dataset in Florida to study high school graduates who entered a four-year public institution straight from high school for the first time in 2000 and 2003, and found that gender differences in STEM attrition were mostly explained by post-enrollment factors such as grades and credits attempted during the first two years of college in introductory courses after controlling for other demographics and pre-college factors. The evidence showed students left STEM due to their lower high school preparation in STEM and consequently lower grades in lower-division STEM courses (Minaya, 2018). Moreover, in the first semester of college career, students who took more than 40% of their courses in STEM were less likely to leave STEM major compared with those who took less than 40% of courses in STEM fields (Olson & Rapporteurs, 2012).

In computing field, George et al. (2022) found two years after the introductory course, over 53% of students indicated an interest in a computing career, but 24.2% students lost interest in computing career two years later compared with the moment of the end of the introduction courses they were in computing majors. The results reported students were more likely to leave the computing career track than at the enrollment stage as 66.6% students initially interested in computing careers maintained their interest, 33.4% lost interest or decided against pursuing a computing career path. This sample shares a consistent pattern of a greater share of men interested in computing careers than female students with the interaction of race/ethnicity and gender. Computing self-efficacy was found a strong predictor of the computing career interest. Comparing different major aspirations between gender in computer science, college freshman women in computer science tended to earn lower high school grades than women in other STEM fields, but higher than did men (Lehman et al., 2016). Lehman et al. (2016) pointed out that freshman female students rated themselves lower in math ability, intellectual self-confidence, than men in computer science. Women in computer science also rated themselves lower academic ability than women in other STEM fields (i.e.g Engineering, Math/Statistics, Physical Sciences). Female students in Computer Science had lower SAT math scores than male students, and women in Engineering and Math/Statistics, but higher than those in Biological Sciences and Physical Sciences. Notably, women in Lehman et al.'s study were more likely than man to be undecided career choices in computer science fields.

In engineering field, Ohland and Lord (2020) studied the role of grades in major selection using a multi-institutional dataset and found engineering students in Chemical Engineering, Civil Engineering, Electrical Engineering, and Mechanical Engineering with lower grades in introductory courses were more likely to leave at various timepoints (i.e., enrolled 3rd semester, matriculated directly remaining 3rd semester, matriculated directly to each major) or changed major. However, they also found the pattern was not the same for students in Industrial Engineering because this discipline was more diverse than other disciplines; even when students had lower grades in introductory courses than the other four majors, students were more likely to succeed in the Industrial Engineering major. Studying college students at a large public university in the US showed reasons why students changed a major were due to lack of sense of belonging in the academic major, or lack of sense of achievements such as negative grades, or a lack of knowledge about the specific field and the careers it offered (Marade & Brinthaupt, 2018). Both students and faculty in the study stated changes in career goals or personal growth and development were very good reasons for changing a major.

Besides academic factors, previous empirical studies have shown monetary and non-monetary factors affecting students' decisions on entering or withdrawing from a STEM college major. The monetary factors include positive perceptions of future earnings (Patnaik et al., 2020; Quadlin, 2020), or ability to pay during college education period (National Academies of Sciences, Engineering, and Medicine, 2016), and conversely, at some institution, higher tuition costs to pursue engineering degrees which often undercuts efforts to increase female and minority students (Stange, 2015). In addition, studies have indicated non-loan aid supports students to earn more credits (Stoddard et al., 2018), but these aids are not sufficient to cover all expenses students spend during colleges. The non-monetary factors affecting students' decisions on entering or withdrawing from a STEM college major include student ability, professional preference, academic preparation (National Academies of Sciences, Engineering, and Medicine, 2016; Zafar, 2013), or contextual factors of institutional type (Chen, 2013), engineering departmental climate (Rincon & George-Jackson, 2016a), and faculty-student interaction (Kim & Sax, 2017). Thus, this study examines financial factor of student SES (i.e. a combination of household income in 2008, father's education, mother's education, father's occupation, mother's occupation), financial aid in terms of merit aid and need aid, as well as non-financial factors of student motivation and self-efficacy in math and science, coupled with college-level academic achievements, collegiate factors and personal conditions to uncover their relationship with choice of a STEM major.

Research problem statement

Despite the increasing demand for workers to occupy positions in STEM fields, federal policy support scheme, the gender and race/ethnicity disparity in STEM major choice may cause disadvantages for underrepresented minority students to realize the economic mobility STEM occupations may provide and prevent the US from innovative competitiveness. Understanding how both high school and postsecondary factors predict the probability of student STEM major choice represent a significant research endeavor and will be the focus of this study.

Addressing the literature gaps

In past studies, using HSLS:09 dataset, researchers have found evidence that student background characteristics (gender, race/ethnicity, SES) and high school experiences contributed to the intention of a STEM major when students reported in their first year of college (Bolds, 2017; Jiang, Simpkins, et al., 2020; Kurban & Cabrera, 2020). Notably, with the same HSLS:2009 dataset, another study added the AP STEM course increases the odds of female students selecting a STEM major more significantly than for male students when students declared their major in the third college year (Jewett, 2019). However, these studies excluded the collegiate supports and personal circumstances when students were in college to consider contextual factors may swing their intention in STEM fields, especially female students in STEM were more likely to change major than male students (Chen, 2013). Furthermore, previous studies have used data from a state-wide university system to study STEM career decision-making process (Castellanos, 2018), or from public universities within a specific US region to study who declared a STEM major (Mau, 2016). There is a call for additional research on the longitudinal prediction of choice actions (e.g. efforts to implement choice goals over time) (Lent & Brown, 2019). The current research addresses this call and extends the literature to examine how collegiate factors and personal circumstances affect the probability of student STEM major choice using a latest national dataset. The dataset allows the researcher to correlate high school freshmen's math and science attainment value, math and science utility value, math and science self-efficacy, college-level STEM credits-earned, STEM grade point average (GPA) to their STEM declared major choice after three years enrolled in college. The overall intent of this study is to inform institutions, programs, and policies which aim to increase the underrepresented students in the future STEM workforce, and to provide appropriate support for students while they are in college, especially students have had disability/special need.

Research questions

The first question examines both high school and college factors in relation to the probability of STEM major choice. The second questions examines only college-level variables in relation to the probability of STEM major choice. Finally, the final question explores what factors predict STEM GPA.

The following research questions guide this dissertation:

(1) To what extent do high school math and science motivation and self-efficacy, collegiate factors, and personal circumstances promote or hinder students' STEM major choice, controlling for student background characteristics?

(2) To what extent do collegiate factors and personal circumstances predict the probability of STEM major choice, controlling for student background characteristics?

(3) What factors predict college STEM GPA?

Significance of the Study

The study will provide a comprehensive understanding what factors mostly explain the likelihood of student major choice over seven years with the aim toward career aspirations in STEM fields. Identifying the explained factors in these models will support institutions, parents, students, and policy makers for initiating appropriate programs to increase women and minorities in the future STEM workforce.

By addressing these research questions, the study will advance previous research in at least four ways. First, previous longitudinal studies have investigated high school math and science self-efficacy, math and science attainment value, math and science utility values in relations to the intention to STEM major when students were in the first-year of college using HSLS:09 data (Bolds, 2017; Itauma, 2019; Kurban & Cabrera, 2020a; Marsh, 2020), or in the second year of college using another longitudinal study ELS:02 data (Wang, 2013). For students who did not declare a STEM major early when enrolling college, they may benefit from acquiring information on and encouragement toward STEM programs during the first year of college to make up their mind (Wolniak, 2016). Moreover, as Wang (2013) explained "some students may still be exploring their major fields of interest during this time frame [sophomore], and others might switch into STEM disciplines later on" (p1112). Unlike previous studies to analyze "planned STEM majors", this study examines *declared STEM majors*. Thus, examining STEM declared major choice is a rationale as students have settled into their major and less likely to change majors at this point in their college career.

Second, empirical studies have examined socialization variables including parents' educational attainment expectations for their child/ren and parents' mathematics and science self-efficacy in helping with math/science homework (Marsh, 2020); parents and teacher's encouragement to study math, parents' discussed STEM programs/articles when students were in high school senior year, and student participation into different math activities such as competition, camps, study group/tutoring, club (Itauma, 2019); parent involvement in student's academic, schooling, career and future plans (Kurban & Cabrera, 2020), in relations to student intention to choose a STEM major. However, socialization, academic and contextual factors at high school differs from the postsecondary level as the content and requirements of undergraduate courses, or projects are more difficult than high school courses, and students' social network may shift to include faculty, peers, and staff from administrative offices such as student services and financial aid to name a couple. In terms of academic performance. Previous studies investigated high school and college academic achievements separately, and few studies examined the direct effect of personal circumstances on STEM major choice. Thus, this study fills these gaps drawing from Situative Expectancy Value Theory (SEVT) and Social Cognitive Career Theory (SCCT) to understand both high school and college academic achievements, environmental supports and personal barriers in student STEM major selection. The SEVT postulates student motivation to achievement-related choices and performance and subjective task values influences their performance. In this study, the subjective task values are math and science attainment value, and math and science utility value. The SCCT assumes students are likely to form interest and choice as they are competent in that domain. When they are confident in math and science, they are more likely to choose STEM major choice for career later.

Finally, there were evidence of differences between gender in declaring a STEM major. Math self-efficacy, postsecondary STEM progress in introductory science lab and advanced math courses are strong predictors of STEM major among community college students (Evans et al., 2020). The ELS:2002 did not measure science self-efficacy (Evans et al., 2020). Thus, this study uses HSLS:09 data which includes science self-efficacy and examines the differences of high school math and science self-efficacy, and STEM college major choice between gender, race, and two groups of four-year enrollees and two-year transferred program enrollees.

<u>Terminology</u>

Collegiate supports in this study include amount of merit-only and need based aid students received at their primary first year institution, research participation with faculty, academic support services, and career support services students used.

Personal circumstances in this study refer students had work schedule interfered with academic performance during attending college, money worry for regular expenses in the calendar year 2015, and disability if students ever had.

Major represents the depth of knowledge of undergraduate curriculum and consists of a set of courses in one or two related fields (Filsecker, 2011).

Credit is a 50-minute-based instruction per week for an academic term assigned to courses or course-equivalent learning (Filsecker, 2011).

Equal opportunities/social mobility: community college

The American higher education system has been shaped by three core beliefs: limited government control, market rationality/competition, and equal opportunities/social mobility through a wider access through a unique structure of community colleges (Filsecker, 2011).

Type of institutions

The federal law distinguishes between higher education institutions (i.e., authorized by the State, accredited by a recognized accrediting agency, and providing at least a bachelor program or at least two year program) and additional higher education institutions (i.e. offering formal instructional programs and a curriculum mainly for students with high school certification) (Filsecker, 2011).

Transfer

There are four types of transfer of student credits from one institution to another: (1) vertical transfer, from a community college to a four-year institution, (2) reverse transfer, from a four-year institution to a community college, (3) lateral transfer, between two four-year institutions, (4) swirling, students enrolled in more than one institutions at the same time (Filsecker, 2011).

STEM

STEM has been previously defined by the National Academy of Engineering and National Research Council in 2009 (National Academies of Sciences, Engineering, and Medicine, 2016). Different definitions of STEM in the US are (a) based on the hard-soft paradigm distinction defined by Biglan (1973), (b) provided by the Classification of Instruction Programs, National Center for Educational Statistics (Department of Education, 2019) that was adapted from the National Science and Mathematics Access to Retain Talent (SMART) Grant (Douglas & Salzman, 2019; Ingels et al., 2015; Jones et al., 2019), and (c) defined by the National Science Foundation (NSF) (Manly et al., 2018).

The present study uses the operational definition provided by the Classification of Instruction Programs, National Center for Educational Statistics (NCES, Department of Education) that was adapted from the SMART Grant (Douglas & Salzman, 2019; Ingels et al., 2015; Jones et al., 2019), which includes Science, Math, and Engineering majors only. The NCES – STEM first-major variable in HSLS:09 data is a binary variable (coded X4RFDGMJSTEM) whether the respondent had declared or decided a STEM or non-STEM major reported during the second follow up interview in 2016.

Table 1. STEM majors defined by NCES

STEM Major	Agriculture, Agriculture Operations, and Related Sciences
	Natural Resources and Conservation
	Computer and Information Sciences and Support Services
	Engineering
	Engineering Technologies/Technicians
	Biological and Biomedical Sciences
	Mathematics and Statistics
	Multi/Interdisciplinary Studies
	Military Technologies and Applied Science
	Physical Sciences
	Science Technologies/Technicians

Student success

This study employs the approach designated in the report of “Barriers and Opportunities for two-year and four-year STEM degrees: systemic change to support students’ diverse pathways” by the National Academies of Sciences, Engineering, and Medicine in 2016. Student success describes students who are interested in STEM majors “are able to make informed decisions about the best course of study for them based on interests, motivation, and career aspirations; understand the variety of potential career pathways the come with STEM degrees; have a clear understanding of STEM content and practices; do not face unreasonable barriers long their pathways that discourage them or make progress impossible; and are aware of connections between STEM and societal issues and concerns.” (NASEM, 2016, p.15).

Math and Science motivation

Math and science motivations refer to the achievement-related motivation of subjective task value, specifically the attainment and utility values in this study. These will be defined in greater detail in Chapter two.

Math and Science self-efficacy

Self-efficacy and confidence are used interchangeably throughout this study. Self-efficacy is a major part of Bandura’s social cognitive model of learning and development that is defined as Bandura’s quote “people’s judgements of their capabilities to organize and execute courses of action required to attain designated types of performances” (Lent & Brown, 2019; Wigfield et al., 2021). The social cognitive theory was extended to apply for career development that Lent & Brown (1994) developed and called it “social cognitive career theory – SCCT”. This study utilizes self-efficacy in math and science courses (Ingels et al., 2014).

Chapter one provides a blueprint of STEM workforce demands in the US, policy to increase women and minorities for the future STEM workforce, college graduate rates in STEM and STEM-related predictive factors, gap of literature review, research questions, significance and limitations of the study. The dissertation is organized as follows: chapter two presents two theories employed in this study and prior literature about the importance of high-school math and science motivation and self-efficacy, and college STEM academic achievements in relations to STEM major choice. Chapter two also illustrates conceptual framework and operationalized variables in this study. Chapter three introduces the data, analytical sample, and analytical strategy. Chapter four presents descriptive statistics and the analytical results in the probability of STEM major choice. Chapter five discusses the interpretations and implications, as well as concludes.

CHAPTER TWO

LITERATURE REVIEW

This study employs two theories, the modern expectancy-value theory (EVT) of achievement choice and performance which is based on *psychological* motivation theory (Eccles et al., 1983; Wigfield et al., 2021; Wigfield & Eccles, 2020), and Social Cognitive Career Theory (SCCT) (Lent et al., 1994), to understand college student motivation, and career preparation in STEM fields. Both theories consider constructs that are domain specific, recognizing variance of students' beliefs and behaviors based on individual self-efficacy, identity, goals, and socializers' (e.g., peers, mentors, teachers, parents) beliefs and behaviors.

The EVT was initially developed to understand gender differences in adolescents' achievement choices and has been labeled recently as the Situative Expectancy-Value Theory (SEVT) due to the situated appropriateness of individual development such as age, gender, and decisions at one point in time and subsequent enactments (Eccles & Wigfield, 2020; Wigfield & Eccles, 2020). The SEVT built on previous expectancy studies. Specifically, research on how expectancies influenced actions was originally from Tolman (1932, 1948), then Lewin (1938) discussed how the value of an activity influenced whether individuals would engage in the activity (Wigfield & Eccles, 2020). In 1957, Atkinson developed further Lewin and Tolman's work by investigating the first formal, mathematical expectancy-value model of achievement motivation. In 1983, Eccles and colleagues extended Atkinson's work by linking "achievement performance, persistence, and choice most proximally to individuals' expectancy-related and task value, and linking in a broader array of more distal psychological, social, situational, and cultural determinants" (Wigfield & Eccles, 2020, p.164).

The subjective task values in the SEVT defined as “personal value a person attaches to tasks being considered for engagement” (Wigfield et al., 2021, p. 91). Values in the SEVT model are subjective because “various individuals assign different values to the same activity; math achievement is valuable to some students but not to others” (Wigfield & Eccles, 2020, p.167). Furthermore, expectancies for success in SEVT are defined student beliefs on how well they will do on tasks either in immediate or long-term future which is similar to Bandura’s construct of self-efficacy (Wigfield et al., 2021).

Together with SEVT to explain psychological motivation, the SCCT - based on Bandura’s Social Cognitive Theory (1977,1986) - comprises of the contextual, person, and *behavior* factors (e.g., self-efficacy, social supports, goal setting) that were assumed to help shape career and educational development across the lifespan (Fouad & Santana, 2017; Lent & Brown, 2019; Wolniak, 2016). The SCCT assumptions and mechanisms based on the triadic reciprocal person-situation interaction model that include personal attributes (i.e. internal cognitive and affective states, and physical attributes); external environmental factors; overt behaviors (i.e. distinct from internal and physical qualities of the person) (Lent et al., 1994). Lent and his colleagues indicated social cognitive mechanisms framework career development comprising of self-efficacy beliefs, outcome expectations, and goal representations. Self-efficacy is derived from four primary sources of learning experiences: mastery or personal performance accomplishments, verbal or social persuasion, vicarious learning (i.e., observation of social models), and physiological and affective is states and reaction (e.g. anxiety associated with task performance) (Sheu et al., 2018). The SCCT postulated math self-efficacy in career choices of women that found college women had lower math self-efficacy than college men, and self-efficacy was the predictor of academic satisfaction and intended persistence across the third and fourth semesters in engineering field (Fouad & Santana, 2017; Lent et al., 2015).

In SEVT, the individual’s self-concept of ability is defined as the extent to which students feel competent in a domain (Jiang, Simpkins, et al., 2020), whereas self-efficacy in SCCT refers to beliefs about one's ability to perform particular academic or career behaviors (Sheu et al., 2018). Both SEVT and SCCT provide a framework that explicitly incorporates gender and race as a person input, but SCCT includes contextual influences at proximal and distal levels in different points in time of career development and educational decision making (Fouad & Santana, 2017; Wolniak, 2016). Over 35 years of SEVT and 25 years of SCCT research on students’ achievement-related motivational beliefs and learning experiences for career development have shown to be successful at explaining students’ math and science motivation and self-efficacy, STEM credits-earned at college, environmental supports and barriers, and STEM major choice for future career (Lent & Brown, 2019; Wigfield & Eccles, 2020).

Situative Expectancy-Value Theory (SEVT)

The expectancy-value theory that Eccles-Parsons and her colleagues framed links achievement performance and choice to individual’s expectancy-related and task-value beliefs. Eccles-Parsons and colleagues (1983) posit that expectancies and values influence performance and task choice directly. Individuals’ expectancies (i.e. domain specific personal efficacy) and values are influenced by their task-specific beliefs such as their self-concepts of ability, perceived task difficulty, and their goals (Wigfield & Eccles, 2020). These beliefs and goals are influenced by their attitudes and perceptions of other people’s expectations for them and their interpretation of self-previous achievement outcomes. Student motivation to achievement-related choices and performance is determined by two constructs: (i) their expectation of success for a given task and valuing of their own competence, and (ii) subjective task value (STV) for a task (i.e., qualities of different achievement tasks and how those tasks influence the individual's desire to do the tasks). Individuals' overall STV is comprised of four components: (i) interest – enjoyment value (i.e., intrinsic value); (ii) attainment value (i.e., how important completing the task is to the sense of self) - the task is central to student see themselves and others see students; (iii) utility value (i.e., whether the individual perceives the task as useful); and (iv) relative cost (i.e., individual's perception of what is lost or given up relative to what is gained when doing any particular task). SEVT has been applied in a number of studies examining students' STEM-related academic major choice intentions and career aspirations (Jiang, Simpkins, et al., 2020; Mills, 2019; Rangel et al., 2020).

This study focuses on two values: attainment value and utility value. Eccles-Parsons et al. defined attainment value as the personal importance of doing well on a given task (Wigfield & Eccles, 2020). Wigfield and Eccles (2020) explained "attainment value derives from the fit of perceived task characteristics with the individual's core self-schema, social and personal identities, and ought selves, that is, the extent to which tasks allow or not the person to the manifest those behaviors that they view as central to their own sense of themselves, or allow them to express or confirm important aspects of self" (p.167). Individual identities can be personal (sense of self as unique) or collective (sense of self as part of a group) (Marsh, 2020). There is evidence that the extent to which students perceive themselves as having a math and science attainment value relates to math achievements, STEM major, and science career (Briggs, 2014; Byars-Winston & Rogers, 2019a; Carlone & Johnson, 2007; Cass et al., 2011; Chemers et al., 2011; Godwin, 2016; Godwin et al., 2013; Itauma, 2019). Math and science attainment values in this study measured students' agreement or disagreement with statements about seeing themselves as a math, science person and others seeing them as a math, science person.

Utility value indicates the extent to which students think math and science are useful to their daily life, college admission, and future jobs. Studies have linked math and science utility to STEM choice outcomes (i.e. course-takings, major choice, and career aspirations) (Gottlieb, 2018; Jiang et al., 2020; Marsh, 2020). STEM course-takings and credits earned reflect the connection of utility value and personal goals because students need math and science for standardized tests such as SAT/ACT to prepare for college admission, or for pursuing their future occupations. The hypothesis holds that attainment value and utility value relate positively to STEM major choice.

Empirical studies using SEVT has identified science self-concept of ability in science among 9th graders was significantly associated with college STEM major choice, STEM high school course-taking, and high school STEM GPA were positively associated with STEM college major in STEM (Jiang et al., 2020). Jiang and her co-authors also used HSLS:09 dataset and focused on subjective task value including intrinsic and utility values. They found math subjective task value did not significantly predict STEM college majors, but science subjective task value was positively associated with STEM major choice. Also, applying SEVT and analyzing HSLS:09 to identify factors that are significantly related to student intention to STEM occupations by the age of 30, Gottlieb (2018) found utility, interest, and attainment values were significantly STEM career plans of White students, but fewer significant relationships between these values and STEM career plans for Black and Hispanic students. Gottlieb also suggested the SEVT constructs are appropriate to model predictions of STEM career plans for students at the bachelor's degree level, rather than for students who may perceive math and science interesting and useful but not planning on earning a bachelor's degree.

Extending the understanding of adolescents' (at ages 13-18) dream jobs and employment realities in the US, a recent empirical study reported the most popular aspirations for females were doctors, veterinarians, teachers, and nurses (Hoff et al., 2022). Hoff and his associates found that doctors were most popular in early adolescents (at ages 13-15), while veterinarians, teachers, and nurses were more popular in late adolescence (at ages 16-18). For males, athletic career aspiration was the most popular aspiration during early adolescence, but become less popular in late adolescent. Other career popular aspirations for male include military, doctors, farmers, and managers. Hoff et al. also explained the variability of all career aspirations decreased from early to late adolescents within the top ten occupations for both females and males, that meant career goals become more varied as more exploration occurred based on Super's life-span model (1980). Females were more likely to aspire to investigative careers compared to males (Hoff et al., 2022), while typical occupations associated with investigative interests include scientists, physicians, geologists, and pharmacists (Wicht et al., 2021).

To support the understanding of student major choice toward their career preparation as there is a strong relationship between college major and post-graduate employment (Huneeus et al., 2021; Patnaik et al., 2020), not only depending on math and science motivation and self-efficacy, but also the environmental factors such as collegiate supports and personal circumstances during the journey of college education, the SCCT framework is incorporated in the study to measure the likelihood of STEM major choice.

<u>Social Cognitive Career Theory (SCCT)</u>

The Social Cognitive Career Theory (SCCT) has five different models (Lent & Brown, 2019), and the present study employs the choice model. The SCCT framework initially consisted of three components: academic and career *interest* development, academic and career *choice-making*, and factors affecting academic and career success (i.e. achievement and persistence) (Lent et al., 1994). Self-efficacy is the central component of SCCT explaining student learning outcomes. Lent et al. posited that students are likely to form an interest in a particular educational domain/activity when they are competent at performing that domain/activity (i.e. they possess self-efficacy, or hold beliefs and confidence in one's ability to succeed in performing the activity) that is also an important source of outcome expectations (e.g., STEM credits earned). Along with self-efficacy and outcome expectations, career-related interests foster particular educational *choice* goals (i.e., intentions to pursue a particular academic major) and take actions that are assumed to be motivated by self-efficacy and outcome expectations (i.e., course-taking). Efficacy research shows how competent students are confident in domains such as math and science, and having math and science interests are more likely to have intention to STEM major (Kurban & Cabrera, 2020). However, in engineering field, Geisinger et al. (2013) reported self-efficacy was also one of six factors contributing to student decisions to leave engineering, besides classroom and academic climate, grades and conceptual understanding, high school preparation, interest and career goals, race and gender. Based on SCCT model, self-efficacy has a direct effect on choice action (i.e., STEM major declaration) and strongly influences ultimate career selection.

Using a meta-analysis approach – including 20% of 196 US and international samples were classified as predominantly racial minority samples, 36% as primarily White American samples, 44% as varied or unknown race/ethnicity, 39% primarily female samples, 32% mostly male samples, and 29% mixed gender samples - to identify predictors of STEM choice, Lent et al. (2018) indicated, the SCCT choice model explained 13% of the variance in self-efficacy, and the negative correlation between environmental supports and barriers was stronger in female than male samples. Particularly, barriers produced a stronger negative path to self-efficacy in male than female samples (Lent et al., 2018). The analysis by race/ethnicity showed the choice model explained 25% in self-efficacy, barriers were negatively predictive of self-efficacy and choice goals with small effects, while supports produced positive and medium paths to self-efficacy with racial/ethnic minority samples.

Numerous studies have revealed math and science self-efficacy have direct effects on student intention to pursue a STEM major and STEM career at age 30 (Bolds, 2017; Jelks & Crain, 2020; Kurban & Cabrera, 2020; Wang, 2013). Using the HSLS:09 dataset, Kurban and Cabrera reported the intention to STEM major was affected most strongly by STEM Readiness comprised of high school STEM GPA, math credits earned and science credits earned (Kurban & Cabrera, 2020). Also, they found math self-efficacy had the strongest impact on math interest as well as science self-efficacy affecting science interest, and math and science self-efficacy had a significant indirect effect on the intention to a major in STEM through math and science interests among high school students. Itauma (2019) utilized HSLS:09 data to study minority female high school students and found a large effect size between mathematics self-efficacy and mathematics identity, science self-efficacy and science identity; a medium effect size between mathematics self-efficacy and mathematics utility; science utility and mathematics utility; science utility and science identity; science self-efficacy and mathematics self-efficacy; and science self-efficacy and science utility**.** In this dissertation, math and science self-efficacy at high school level are hypothesized to associate positively with students' STEM major choice. Math and science self-efficacy are measured by student expectancy in math and science tasks, specifically in tests, assignments, textbook understandings, and mastery skills. Based on student excellence on these tasks, they reached mastery level as one of four learning self-efficacy measure scales in SCCT.

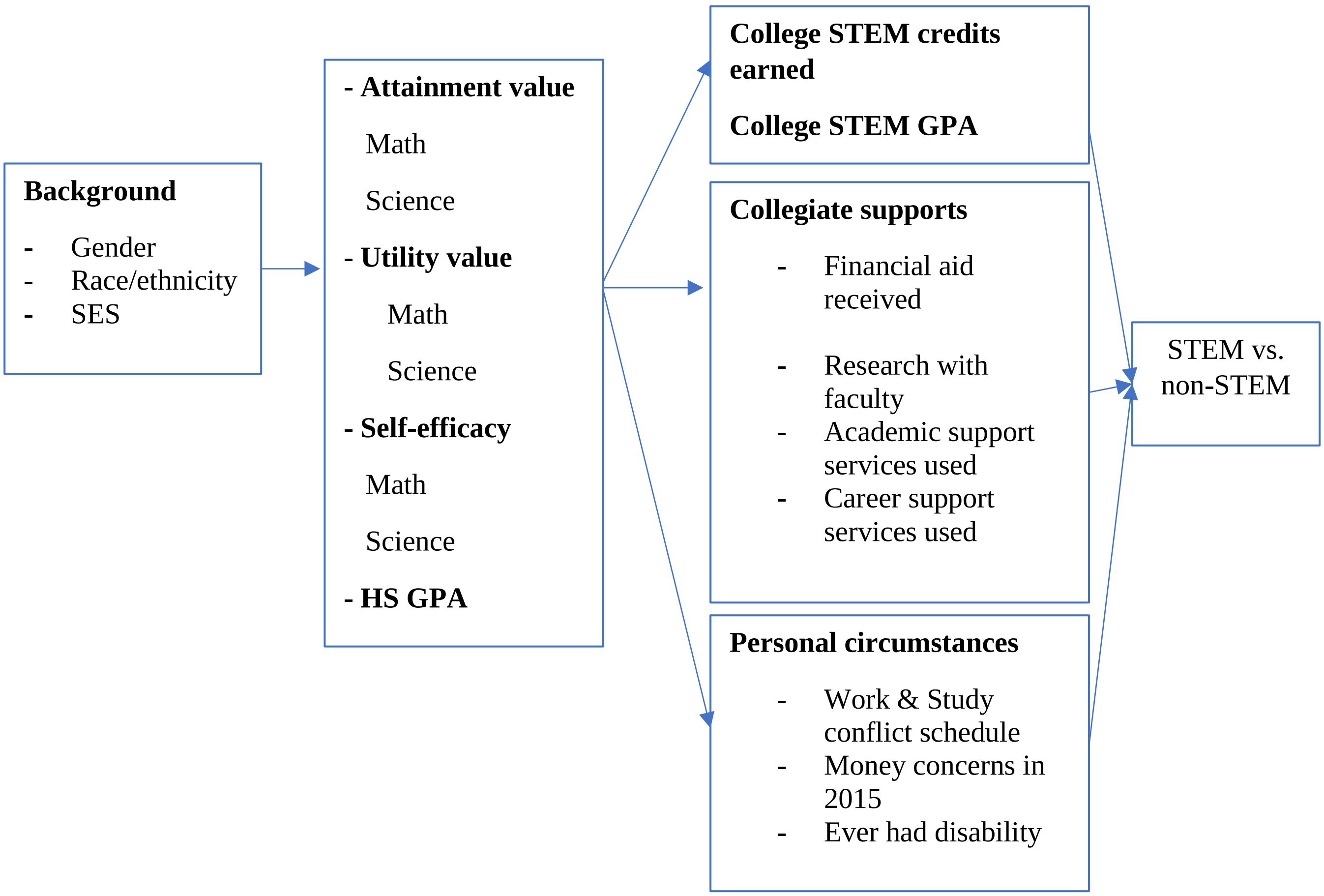
As explained early, SCCT choice model includes contextual supports and barriers together with self-efficacy and outcome expectations. The c*ontextual supports* are assumed to promote choice goals and actions, while *barriers* discourage goals and actions respectively. Environmental support factors include parent's approving of the major (Zafar, 2013), parental involvement (Kurban & Cabrera, 2020), college prep program offered at high school such as Mathematics Engineering Science Achievement (MESA) (Alvarado & Muniz, 2018) that are predictors of STEM major choice. This study extended previous literature by examining collegiate support factors including merit-aid and need-aid students received in their primary first year institution; research participation with faculty, academic support services students used (e.g. writing center mentoring), and career support services such as job placement.

In terms of barriers to entrance into a STEM field of study, Wang (2013) reported number of weekly work hours, receiving remediation, and being enrolled full-time did not have any significant effect. None of these barriers differ significantly across racial, gender, and social economic status groups (Wang, 2013). Previous studies showed financial aid as a support resources for students to retain and graduate a STEM degree (Schwintz, 2019; Wolniak, 2016), but to my knowledge, no prior study have examined specifically a contextual barrier in a specific calendar year prior to a third-year college education on how money worry for regular expenses affect students' STEM major choice. This study will fill this gap.

The conceptual framework of this study is drawn in the Figure 1 below.

Conceptual Framework

Figure 1. Conceptual Framework



Empirical studies on gender differences in STEM major choice

Extant research has investigated gender differences in STEM major choice based on influences such as monetary considerations, STEM climate, STEM disciplines, academic achievements, and personal preferences. Research has shown that these factors operate through push and pull mechanisms that either attract or detract differential patters of gender participation in STEM. A huge literature showed evidence of gender differences in STEM major choice. There is a wide gender gap of bachelor's graduates in STEM majors between 36% for females and 64%

for males in 2015-2016 (US Department of Education, 2016). Ample research on determinants of major choice have uncovered the effect of expected earnings and non-monetary factors such as student preferences and student ability (Jiang, Link to external site, et al., 2020; Patnaik et al., 2020; Pu et al., 2021; Zafar, 2013), or college major traits (Ganley et al., 2018). A study showed no evidence of gender differences in ability to major in science and engineering, whereas the gender disparity exists in preferences to choose a major (Zafar, 2013). Specifically, the gap between female and male students in engineering is 60% due to differences in preferences, while this gap is 30% in student beliefs in enjoying to study engineering. Using The Multiple-Institution Database for Investigating Engineering Longitudinal Development (MIDFIELD) data of 11 institutions (1987-2014) compared with the data 2013 of American Society for Engineering Education (ASEE – 349 institutions), Orr et al. (2020) indicated that White male students were dominant in all five top disciplines of Mechanical Engineering, Civil Engineering, Electrical Engineering, Chemical Engineering, and Industrial/Manufacturing/Systems Engineering (Orr et al., 2020). These five majors accounted for 75% of engineering graduates in MIDFIELD data and 61% of ASEE data. Male student enrollment was dominant in Mechanical and Electrical Engineering. In Industrial/Manufacturing/Systems Engineering, more White female enrollees in MIDFIELD data than in ASEE data. The researcher acknowledge the trend of higher female college enrollment rate at US campus compared with male peers recently (Tran et al., 2021).

Gender differences in STEM major choice and career aspirations have been examined under various perspectives such as sociology, psychology, economics, and education (Cheryan et al., 2017; Jiang, Simpkins, et al., 2020; Patnaik et al., 2020; Quadlin, 2020; M.-T. Wang & Degol, 2017; Zafar, 2013). Economically, male students are tied to choose major which significantly higher future earnings than women's major decision when both genders have the same major preferences (Quadlin, 2020), particularly STEM graduate earnings are higher than non-STEM graduates (Huneeus et al., 2021). Men are more likely to major in higher earning STEM and business fields, while women are more likely to lower paying fields such as education, humanities, and social sciences (Arcidiacono et al., 2013; Quadlin, 2020). However, the effect of expected earnings on major choice are unlikely to "provide much guidance to the likely effects of differential pricing by program" indicated by Stange (2015). Stange investigated higher prices (i.e. differential pricing) for engineering, business, and nursing at 50 large US public research universities from 1990 to 2010. Stange's study results showed higher tuition for engineering appears to undercut efforts to increase female and minority representation in engineering (Stange, 2015).

Much of the sociology, psychology, and education research indicated that some majors have more dominance of male college students than female students, particularly larger male participation in computer science, engineering, and physics majors than in biology, chemistry, and mathematics (Cheryan et al., 2017; Lehman et al., 2016). Cheryan at al. (2017) explained the gender disparity was due to (a) masculine cultures that signal a lower sense of belonging to women than men, (b) a lack of sufficient early experience with these fields, and (c) gender gaps in self-efficacy. Furthermore, Wang and Degol (2017) summarized six explanations for women's underrepresentation in math-intensive STEM fields in the US over three decades including (i) cognitive ability on standardized tests such as SAT/ACT, (ii) relative cognitive strengths, (iii) occupational interests or preferences, (iv) lifestyle values or work-family balance preferences, (v) field specific ability beliefs, and (vi) gender-related stereotypes and biases. According to Wang and Degol, women earned majority of bachelor's and advanced degrees in biological/biomedical sciences, environmental engineering, rather than in math, computer science, physics, and mechanical or electrical engineering. This gender disparity was found consistently when high school course-takings were analyzed using three longitudinal studies conducted by NCES (Riegle-Crumb et al., 2012). Riegle-Crumb et al. reported high school senior male and female students were equally likely to take calculus course, but boys had higher likelihood to take physics, and this difference was larger when students selected a STEM major choice after two years of high school completion. This finding was notable to "undermines the assertion that women's underrepresentation in physical sciences, engineering, math, and computer science fields was due to deficits in prior achievement" (Riegle-Crumb et al., 2012, p.1066). Male students had on average the higher probability to select a major in physical sciences, engineering, math, and computer science than female peers across three national longitudinal datasets (Riegle-Crumb et al., 2012).

Among college students, female STEM entrants were more likely to switch major than their male counterparts (Chen, 2013; Minaya, 2018), and female freshman students were less likely to choose a STEM major than male peers at four-year institutions (Moakler Jr & Kim, 2014) or regional public institutions (Mau, 2016). Gifted students in Nebraska with the odds ratio of STEM majors for males being 5.124 times than of females (Vu et al., 2019), while for first-generation college students, female students are less likely to choose male-dominated majors (Wright, 2019). In other words, male students were more likely to stay in STEM majors compared to female students (Fouad & Santana, 2017; Park et al., 2020; Sheu et al., 2018). Interestingly, female students took more college advanced courses in all major categories except STEM they took fewer, and high school GPA predicted advanced course enrollment in college (Shewach et al., 2019). Using mixed methods to analyze first-year students from the North Carolina Roots of STEM Success Dataset, Stearns and her colleagues reported female students had higher STEM GPA than male students (Stearns et al., 2020). Furthermore, the higher STEM GPA in the first year, the higher likelihood students declared majors in biology or STEM. However, having higher non-STEM grades in the first year may pull students to non-STEM major. When using relative advantage of STEM grades vs. non-STEM grades, Stearns et al. found that the difference in STEM and non-STEM grades increased, the difference by gender in declaring a STEM major is reduced (Stearns et al., 2020).

Empirical studies on race/ethnicity differences in STEM major choice

Many empirical studies have also reported the differences in STEM major choice among race/ethnicity. Based on a synthesis research of minority students' participation in college STEM fields, the disproportionality of racially minoritized students is related to pre-college preparation, systematic school segregation, reduced levels of psychological factors associated with STEM success, lower levels of family social, cultural, and financial capital supporting student academic outcomes (Bottia et al., 2021). Typically, STEM major choice was dominantly White male students (Riegle-Crumb et al., 2012; Thompson, 2020). Asian and Black students tended to be more likely to choose STEM majors than White students in the ELS:2002 data (Li, 2019), but adding high school GPA, AP scores, SAT scores in the hierarchical generalized linear nested model, the statistically significant difference between Asian and White students was eliminated, and adding school variable resulted no difference between Black and White students because Black students tended to come from more racially diverse high schools. Lehman (2016) reported Asian American/Asian women, who are not defined by the NSF (2015) as an underrepresented minority group, have increased slightly their participation in computer science, while Latinas' and African-American women's representation in computer science decreased (Lehman et al., 2016). Also using ELS:2002 data, Nix and Perez-Felkner applied intersectionality analysis and reported the known pattern of less likelihood women declared majors in STEM fields than men (Nix & Perez-

Felkner, 2019). The study indicated Black students had similar higher probabilities to declare a STEM majors than White peers as Li (2019) reported. Nix and Perez-Felkner analyzed math difficulty orientation and found the higher difficulty orientation, the higher probability of STEM major choice students had.

African-American and Hispanic men in undergraduate engineering program reported higher average self-efficacy in science, math, and engineering courses than White men after controlling for personal, environmental, and behavioral factors (Litzler et al., 2014). Similarly, Moakler and Kim (2014) studied first-time, full-time, US freshman students and found African American and Latina/o students were equally as likely to choose a STEM major as were White or Asian American students. Their findings reported 23% of the matriculating minorities were pursuing a STEM major. Another longitudinal restricted-use BPS:04/09 data for predicting STEM major in the first year of college indicated that Black students were more likely to select a STEM major than White and Asian students (Wolniak, 2016). Moreover, a recent study investigated the differences of high-school calculus enrollment among White, Black, and Hispanic students and found math self-efficacy beliefs were not positively associated with enrollment for any group (Thompson, 2020), notably a negative association for Hispanic males. Furthermore, Thompson (2020) reported the math utility value was associated with enrollment for only Hispanic females but no other groups, and math attainment value predicted calculus enrollment for Black males. This may explain why there is little gap in terms of math self-efficacy among race and ethnicity, but the perceived values in math vary further differ STEM major choice among racial/ethnical students at college level. Thompson (2020) indicated that more research is needed on other underrepresented populations, including persons with disabilities and American Indians, Alaskan Natives, Native Hawaiians, and

Pacific Islanders, so this study includes these groups of students in comparison to White students and non-disable/special needs students.

Two-year versus four-year enrollment pathways toward a STEM bachelor's degree

The path of college enrollment to obtain a bachelor's degree has changed remarkably over the last three decades instead of the norm of high school graduates enrolling in a baccalaureate program and earning a degree in four years (National Academies of Sciences, Engineering, and Medicine, 2016). The college student population has witnessed the increase of minorities to two-year colleges and many recent state policies expressly support transfer students from two-year colleges to four-year institutions (National Academies of Sciences, Engineering, and Medicine, 2016). A national representative sample of 2012/14 Beginning Postsecondary Students Longitudinal Study indicated that one-third of students enrolled in the bachelor's degree program changed their major within the first three years of college enrollment compared with 28% rate of students enrolled in the associate's degree program (Department of Education, 2017). The report also provided 35% of students who had declared a STEM major had changed their field of studies within three years. Students in natural sciences or mathematics had higher major change rate than students in other STEM fields such as computer science and engineering. Previous qualitative research suggested that engineering faculty played an important role for successfully transferred students from community colleges to a four-year institution because they encouraged and supported students scholar abilities and professional development (Zhang & Ozuna, 2015). Students gained more academic confidence at community colleges, affordable tuitions and flexible schedules to explore major choice such as foundation classes, and strong interactions with faculty before transferring. Moreover, community colleges are popular destinations for transferring students due to lower cost, increased accessibility, proximity to students' homes relative to four-year institutions (National Academies of Sciences, Engineering, and Medicine, 2016). The sample of this study includes students enrolled into a primary two-year institution then transferred to a four-year colleges and universities compared with students straightly enrolled at a four-year college from high school.

The following section outlines operationalized variables for the study.

Constructs

Outcome variable

STEM majors are defined by NCES categorizations. This binary variable refers to student response (Yes/No) to the HSLS:09 second follow-up survey in 2016 for their reference undergraduate degree- first major field of study.

Predictors

High school-staged factors in this study include student background characteristics such as gender, race/ethnicity, social economic status (SES); math and science attainment value, math and science utility value, math and science self-efficacy in 9th grade; high school GPA.

Postsecondary factors describe STEM credits earned and STEM GPA during the first three years of college, collegiate supports, and personal circumstances.

High school math and science attainment value, utility value, and self-efficacy

A vast literature have examined high school academic achievements (i.e. high school GPA), math and science motivations are associated with intentions to enroll a college STEM major (Itauma, 2019; Jiang et al., 2020; Kurban & Cabrera, 2020; Tran et al., 2021; Vu et al., 2019; Wang, 2013). Jiang et al. (2020) utilized HSLS:09 to measure math and science motivations in predicting STEM courses-taking and STEM GPA at high school, and STEM major at college. Jiang et al. found that female students had lower math and science subject task values, and were less likely to enroll in STEM majors at college than male peers, although they gained higher grades than male students in STEM courses at high school. Math self-efficacy is a strong predictor of STEM major choice based on a national freshman dataset of Higher Education Research Institute (Moakler & Kim, 2014). All math and science attaintment value, math and science utility value, math and science self-efficacy variables are composite scores that were standardized to a mean of 0 and a standard deviation of 1, with the coefficient of reliability for the scale is .65 or higher.

The attainment value is measured by two items "You see yourself as a math person" and "Others see you as a math person" with the response Likert scale of 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree.

The utility value is measured by three items: "Math is useful for everyday life", "Math will be useful for college", "Math will be useful for a future career". The response scale is 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree. Similar items and scales apply for science.

Math self-efficacy was developed using four items, which included, "You are confident that you can do an excellent job on tests in this course;" "You are certain that you can understand the most difficult material presented in the textbook used in this course;" "You are certain that you can master the skills being taught in this course;" and "You are confident that you can do an excellent job on assignments in this course". Respondents reported on a scale of agreement or disagreement: 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree.

Similarly, a set of science attainment value, science utility value, and science self-efficacy were surveyed the same items and scales of math subject.

High school GPA

There is a strong link between high school math and science achievements and the declaration of a postsecondary education major (Riegle-Crumb et al., 2012) when considering pre-college factors. High school GPA is a strong predictor of college match as academic merit factors (Lee et al., 2017; Lent et al., 2018; Vu et al., 2019), persistence among first-time students (Stewart et al., 2015), college enrollment, college graduation (Allensworth & Clark, 2019).

Collegiate factors

College academic achievements: STEM credits earned and STEM GPA

Knowledge development in math and science are continuum in the pathway of high school to college education. Earning credits in introductory science lab credits, and calculus or advanced math credits showed positive association with the odds of a community college student choosing a STEM major (Evans et al., 2020). Moreover, majors in computing and engineering typically have more required courses than majors in the social sciences and humanities (Lee et al., 2021). The higher math ability, the more likely college students enrolled in STEM major (Wang & Degol, 2017). The first-year STEM-related coursework is particularly important in helping students declare STEM majors at four-year institutions (Wolniak, 2016). When students were in the first-year college, the higher STEM-related grades they gained, the higher probability they declared in biology and/or physical science/engineering major (Stearns et al., 2020). This study uses STEM credits students earned and STEM GPA from their transcripts in 2013-2016 to demonstrate their academic achievements.

Extensive research has been done on the effect of substantial financial resources, educational support, and social support to engage and develop students' STEM focus, particularly for female and minority students (Glennie et al., 2019; Kim & Sax, 2017; Moakler Jr & Kim, 2014; Oseguera et al., 2020; Pu et al., 2021; Schwintz, 2019; Wolniak, 2016). College supports in this study include merit aid, need aid, undergraduate research with faculty, academic and career support services on campus students used.

Financial aid

Students in the US receive financial aid (merit- and need-based) from federal, state, colleges, communities, and private donors for their college education. The merit-based, need-based scholarships, academic scholarships, and athletic scholarships are popular depending on student ability and background. The effects of nonloan aid is significantly associated with higher rates of college graduation and gives students flexibility to focus on studies and less time to work (Stoddard et al., 2018). Using Montana University System dataset that contained individual-level data from secondary school through college, Stoddard et al. (2018) found that nonloan aid was associated with positive academic outcomes: higher GPAs, enrollment in more credits, a greater probability of majoring in a STEM field, and being more likely to be retained in the next year (Stoddard et al., 2018). Investigating the effect of student loans on academic outcomes, Stoddard and colleagues' (2018) findings suggested that students who had access to student loan debt took more credits, but as the size of loan increased, students' academic performance decreased (Stoddard et al., 2018).

At institutional level, Ohio State University's new president announced a 10-year, $750 million investment in research and a plan to make undergraduate education debt-free (Flaherty, 2021). Some evidence from a study of students pursuing engineering degrees at two public research universities suggested that students were less likely to choose an engineering degree if institutions charged differential tuition, particularly for low-income students (George-Jackson et al., 2012), and cost was also a disadvantage of changing academic majors (Marade & Brinthaupt, 2018). Furthermore, first-year college students with low academic standing and GPA changed their academic majors in order to keep their financial aid or scholarships and rescued what they can of their education.

Accessing to financial aid policies and procedures are particularly at the disadvantage, or "hidden inequality", for students with disabilities at large public four-year institutions that may affect the enrollment and graduation rates, especially low-income and first-generation students with disabilities (Perlow et al., 2021). The present study examines the effects of merit aid and need aid as a source of collegiate supports on the probability of a STEM college major choice.

Collegiate support: Undergraduate research with faculty

Research on student – faculty interaction has reported a positive association with student outcomes such as college GPA (Chen, 2013); cognitive outcomes (e.g. the ability to understand science and technology), affective outcomes (e.g. positive self-concept); civic outcomes; spiritual outcomes, and vocational outcomes (Kim & Sax, 2017). However, the interactions between faculty and students varies by student background characteristics as Black students interacted more frequently with faculty for course-related matters than other racial groups, and they were less likely to assist faculty with research and most likely to report experiencing racial/ethnic discrimination (Park et al., 2020). Park et al.'s study (2020) reported female, Black, and Latinx students were more likely to leave STEM by the fourth year of college than male, White, and Asian American peers. Notably, Black students were more likely to participate in research when they attended institutions that offered undergraduate research experiences to first-year students as part of a structured program, and they were more likely to participate in undergraduate research programs than White students (Figueroa et al., 2013). Figueroa and her colleagues also found students at private institutions were more likely to participate in research programs than students at public colleges and universities. With the increasing diverse student population on campus, undergraduate research participation increased students' likelihood of retentions in STEM majors, but less likelihood for sexual minority students – LGBQ – than heterosexual counterparts (Hughes, 2018). The single-item measurement whether students participated in research with faculty in this study will provide explanation of research experiences students advance their knowledge and skills because scientific research is prerequisite to pursue in STEM fields either for a major or a future career. The single item of participation in research is not possible to capture the quality of that experience.

Collegiate support: Academic and career services

Academic and career supports are critical for high school students transition to postsecondary education, and from colleges to workforce. Academic and career support services include mentoring, tutoring, writing center, career planning, job placement. As Wolniak (2016) suggested the information students received during the first-year of college could make them change their mind about pursuing a STEM major among students did not have intention in STEM majors when they enrolled at college. The combination of encouraging students to attempt 15 or more credit hours per semester and providing enhanced advising appears to have helped students make greater progress toward graduation (Erwin et al., 2021). Although these supports are available at institutions, the usage of contextual supports for the effectiveness of academic, social, and career preparation varies by gender and races, particularly at departmental level. Studying undergraduate students from nine institutions in the US about their department climate, Rincon & George-Jackson found Black female engineering students experienced a more negative climate than their non-Black female peers, and students perceived negative institutional climate reduced their department climate scores (Rincon & George-Jackson, 2016). There are many STEM intervention programs, especially when students were in transition from high school graduation to college enrollment, to provide social and academic supports for students to be familiar and engaged in STEM studies such as academic advising, exposure to STEM, structured learning and tutoring, hands on experience and research, mentoring and networking (Ashley et al., 2017; Rincon & George-Jackson, 2016). These supports have showed student retention in STEM programs (Oseguera et al., 2020).

The academic support services questioned students whether they used academic support services (for example, tutoring or writing centers). This is a binary variable in this study.

The career support services questioned students whether they used career planning or job placement services or not. This is a binary variable in this study.

Personal Circumstances

Work schedule interfered with academic performance during college

Research on working during college have provided contradictory findings. Sixty two percentages of US undergraduates work for pay while enrolled in college (Douglas & Attewell, 2019). On the one side, benefits of working during college were earnings for college education, food and housing expenses, and students who experienced working during college had higher earnings after graduation (Douglas & Attewell, 2019). On the other side, working during college cause lower academic achievements, retention and graduation rates, and longer time-to-degree because they had to split study and work time (Darolia, 2014). Chang et al. (2014) analyzed repeated-measures freshman and senior student surveys and found working full-time while attending school was negatively associated with underrepresented students' chances of persisting in a STEM field (Chang et al., 2014). However, Wang (2013) reported number of weekly work hours did not have any significant effect on STEM entrance. The current study aims to test whether student work schedule interfered with academic performance during college negatively associates with STEM major choice as a personal circumstance. The survey respondents reported their agreement or disagreement on the survey based on a Likert scale: 1=Strongly agree, 2=Agree, 3=Disagree, 4=Strongly disagree.

Money worry for regular expenses in 2015

Coupled with the "work schedule interfered with academic performance" variable, this variable refers to student worry of money in a specific calendar year of 2015. This study extend literature by taking into account student's specific money worry in a calendar year after enrolled into college to help identify whether this barrier variable lowers the probability of STEM major choice.

Disability/Special Needs

There are 10.6% of US population in the age of 16 – 64 with disabilities in 2016 (National Science Foundation, 2019). Disabilities are hearing difficulty, vision difficulty, cognitive difficulty, self-care difficulty, ambulatory difficulty, and independent living difficulty. The number of students with disabilities is increasing on campus and the estimated proportions of higher education students with disabilities are between 11-12% (Kimball et al., 2016). Using HSLS:09 dataset, Bittinger studied specific group of students with disabilities in predicting their STEM major choice and reported female students had lower odds ratio than males in terms of the likelihood of pursuing a STEM major (Bittinger, 2018). Bittinger's study used STEM outcome variable as student intention to a STEM major upon enrollment in the first-year college. As indicated earlier, the current study explains STEM major choice when students were in their third-year college to overcome the previous limitations of freshman or sophomore stages Wang (2013) suggested.

Perlow et al. (2021) analyzed web-based financial aid information of 51 public four-year institutions in the US and reported the hidden inequality for college students with disabilities. Students with disabilities had to bear more higher education cost of living than their peers such as adaptive equipment, and at some institutions, these students had to pay additional fee-based academic services, e.g. note-taking, one-on-one tutoring (Perlow et al., 2021). In terms of financial aid adjustment policy that students had to advocate themselves, especially Obama era professional judgement guidance was rescinded in June 2020, nearly 40% of the website financial aid information did not reference disability in any way and fewer than one in five institutions provided clear steps for students to engage this policy. Furthermore, only 20% of institutional websites provided information of time limits to degree completion as federal aid allows 150% of the time allotted for degree completion. This variable is hypothesized to be negatively related with STEM major choice given student background (gender, race, SES) and received financial aid.

Types of disabilities in HSLS:09 were defined if students ever: 1) had a serious difficulty concentrating, remembering, or making decisions, 2) had been told by a health or education professional that he/she had ADHD or ADD (Attention Deficit Hyperactivity Disorder or Attention Deficit Disorder), 3) had a learning disability, 4) was deaf or had a serious difficulty hearing, 5) was blind or had a serious difficulty seeing, or 6) had any other disability or special need.

Control variables

Gender

The public HSLS:09 dataset provides binary gender variable, male or female.

Race

There are eight groups of the original race/ethnicity variable in HSLS:09 dataset. The racial categories are based on the Census definition. This variable is recoded into five groups as detailed in Chapter three.

Social economic status

An enormous literature has used social economic status (SES) including parent education, parent income to study the probability of STEM major choice. Students whose parent occupations in STEM fields were approximately 1.6 times more likely to choose a STEM major (Moakler Jr & Kim, 2014), and 4.23% higher likelihood of being retained in STEM (Hughes, 2018). Lichtenberger and George-Jackson utilized a statewide dataset and found low-income students were more likely to be interested in STEM majors than higher income students (Lichtenberger & George-Jackson, 2012). However, Moakler and Kim (2013) reported students whose parents' income was equal to or greater than $250,000 were approximately 1.7 times less likely to choose a STEM major. The family SES does not predict the likelihood of STEM enrollment by itself, but it interacts with other predictors such as gender, race, and math achievement (Niu, 2017). Using ELS:2002 longitudinal data, Niu (2017) found family SES promoted the STEM enrollment for female and Black students. Students from lower SES often lack math preparation, information about STEM majors and occupations to make their major choice. Thus, low-income students were less likely to pursue STEM majors (Niu, 2017) but more likely to enroll at four-/two-year for-profit, or community colleges, while middle-/high-income counterparts were more likely to attend four-year institutions and selective colleges (Dynarski et al., 2018; Fry & Cilluffo, 2019; Oseguera & Hwang, 2014).

Chapter two synthesized theoretical frameworks of SEVT and SCCT employed for this study to test the STEM major choice students declared in their third-year college of the HSLS:09 data.

The literature review illustrated the diverse studies with various datasets to demonstrate how math and science motivational beliefs, math and science self-efficacy, STEM credits earned at college, STEM GPA, and environment factors were related to STEM major declaration. This chapter also provided gaps of literature that the present study will address. The next chapter will present the research design, data source, analytical sample, and data analysis.

CHAPTER THREE

METHODOLOGY

In this chapter, four sections are organized as follows. The first section provides the rationale for the correlational survey research design and re-states the research questions. The second section details the sampling design and analytical sample used in the research. The third section describes the theoretical constructs and how these are measured as variables within the dataset. The last section outlines the statistical procedure employed to answer the research questions.

Research design

Administered by the National Center of Education Statistics, The HSLS:09 dataset is the most up-to-date national longitudinal dataset and the most desirable dataset from which to explore the choice of STEM major. The HSLS:09 survey not only explicitly asked students about their declaration of major in third college year, but also provided extensive information about course taking history (e.g., STEM related classes) and contextual factors in college. Moreover, HLS:09 permits the researcher to explore to what extent 9th graders' math and science attainment values, math and science utility values, math and science self-efficacy, and college-STEM credits earned in college as well as college-STEM GPA consistently predict students' STEM major choice over a period of seven years. Furthermore, the HSLS:09 provides variables to investigate how collegiate factors and personal circumstances affect the probability of declaring a STEM major. To my best knowledge, no study has used the third-year STEM major declaration as an outcome variable in relation to both high-school and postsecondary education periods, collegiate factors, and personal circumstances.

The correlational survey research is appropriate to explain the relationships between high school and college academic achievements, individual achievement-related motivations, and environmental factors because of the following reasons. First, student development occurs over time, so observing student developmental growth in the process of exploring one's attainment, ability, and considering educational and occupational paths from high-school age through college is critical to understand emerging adulthood (Jiang et al., 2020). Second, the large random sample of US students provides high reliability for representatives from diverse background to test theories, rather than an exploratory qualitative approach with in-depth insights of a small sample in terms of transferability. The motivational and career behavior constructs under the SEVT and SCCT have been tested for decades reflecting the validity and reliability of the constructs to serve this study. Third, time and costs to collect such national representativeness for the research purpose are only feasible with the resources of secondary dataset (Pérez-Sindín, 2017). HSLS:09 public dataset permits the feasibility of the current research study, which examines the following questions:

<u>Research questions</u>

(1) To what extent do high school math and science motivation and self-efficacy, collegiate factors, and personal circumstances promote or hinder students' STEM major choice, controlling for student background characteristics?

(2) To what extent do collegiate factors and personal circumstances predict the probability of STEM major choice, controlling for student background characteristics?

(3) What factors predict college STEM GPA?

<u>Sampling design</u>

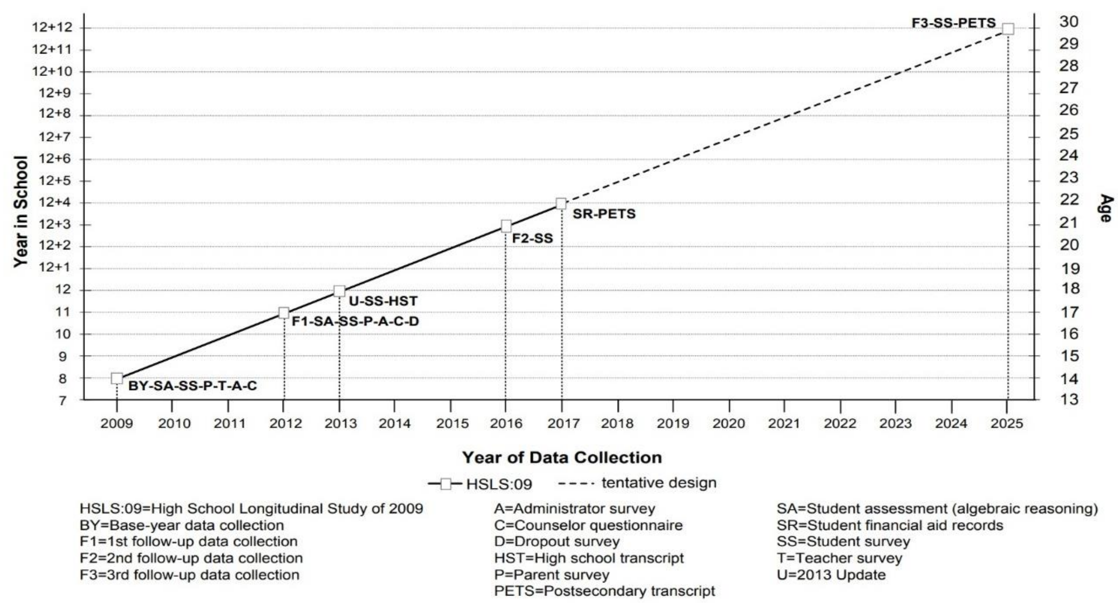
Respondents represented in the HSLS:09 dataset were sampled in 2009 from among ninth-graders in the US using a two-stage process:

First, the survey population for the full-scale study of HSLS:09 consists of all ninth graders in the 50 states and District of Columbia enrolled in *regular public schools,* including state department of education schools that include 9th and 11th grades; *and Catholic and other private schools* that have 9th and 11th grades. Excluded for HSLS:09 are the following: schools with no 9th or 11th grade; ungraded schools; Bureau of Indian Affairs schools; special education schools; area vocational schools that do not enroll students directly; Department of Defense schools; and closed public schools (Ingels et al., 2011, p.153). Among 1,889 eligible schools, 944 schools participated in the HSLS:09 survey. The strata were cross-classifying state with school type (i.e public, Catholic, private), and four-level urbanicity (i.e. urban, suburban, town, and rural based a Census definitions).

Second, 25,206 eligible students were randomly sampled from school ninth-grade enrollment lists (about 27 per school) and 21,444 students participated, accounting for 86% weighted in 2009. HSLS:09 school and student samples are nationally representative and also state representative for a subset of 10 states (Ingels et al., 2011). In the second data wave (the first follow up in 2012), there were 25,184 eligible students and 20,594 students participated that reached 82% weighted. In the third data wave (2013 Updates and High school Transcripts), which occurred after the majority of students in the sample finished high school, the weighted percent was 73% for 25,168 eligible students and 18,558 participated students. In the second follow up data wave in 2016, 25,123 eligible students and 17,335 participated students. The second follow-up was conducted between March 2016 and January 2017 (hereafter called 2016 wave) three years after the majority finished high school. As such, the 2016 wave collected information on college experience (i.e., major choice, retention, academic performance). Figure 1 presents the timeline employed in the HSLS:09 data collection, and the base-year survey conceptual map (Ingels et al., 2011).

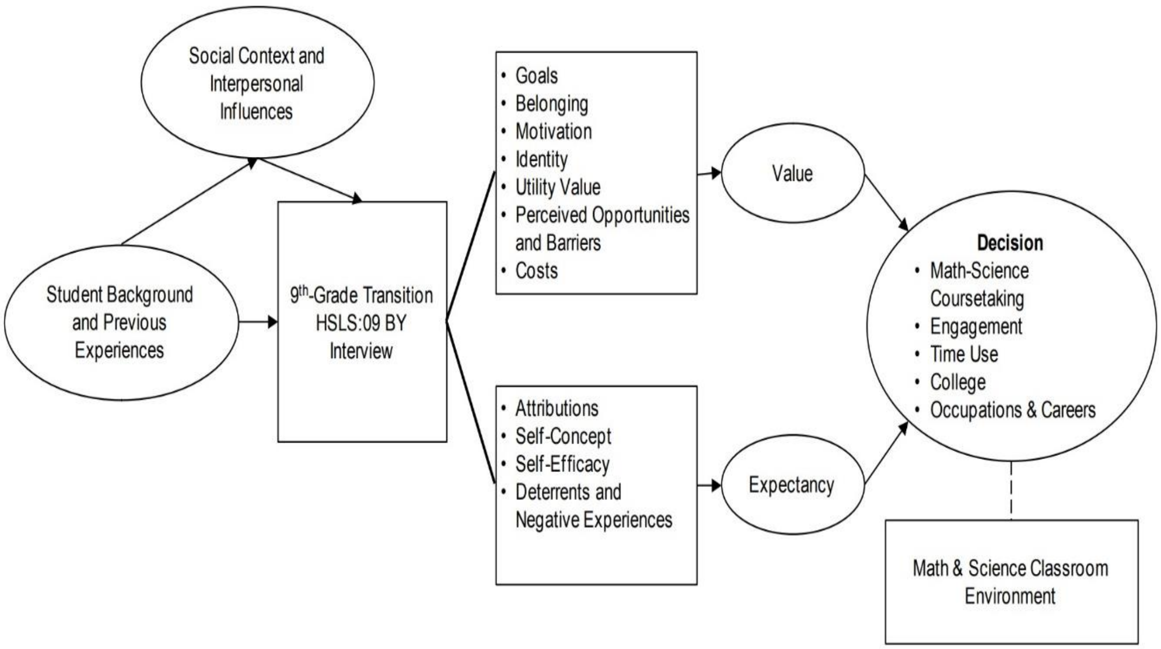
The total sample of the HSLS:09 public dataset is 23,503 students.

Figure 2. HSLS:09 data collection timeline



SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09).

Figure 3. HSLS:09 base-year student survey conceptual map



Variables

Based on the conceptual framework of this study in chapter 2, this section explains in detail variables utilized in this study, each variable scale, and how variables are coded for analytical purposes (*see* Appendix A).

Outcome variable

STEM major choice. Students responded to the second-follow-up data collection in 2016 and this variable was created by NCES. The binary variable indicates 1 for majors students declared in STEM fields and 0 for major choices of non-STEM. This variable is based on the "first major" (i.e. the major students reported first) for the undergraduate degree major field of study (Duprey et al., 2018).

Table 1. STEM majors defined by NCES

STEM Major	Agriculture, Agriculture Operations, and Related Sciences
	Natural Resources and Conservation
	Computer and Information Sciences and Support Services
	Engineering
	Engineering Technologies/Technicians
	Biological and Biomedical Sciences
	Mathematics and Statistics
	Multi/Interdisciplinary Studies
	Military Technologies and Applied Science
	Physical Sciences
	Science Technologies/Technicians

Control variables

Gender is a binary variable, coded (0) = male, and (1) = female students. It is important to note that students were identified as either male or female, therefore the assignment may not reflect students who do not self-identify in this way. Respondent's gender was taken from the base year of 2009. If the gender indicated any inconsistency among 2009 student questionnaire, parent questionnaire, and school-provided sampling roster, X1SEX was coded based on manual review of the sample member's first name (Ingles et al., 2011). Among 23,503 total sample of the public HSLS:09 dataset, only six students had missing value.

Race/ethnicity is a categorical variable with six levels, of which White is the reference group. Based on student questionnaire, the following race/ethnicity groups were created by NCES: (1) America Indian/Alaska Native, non-Hispanic, (2) Asian, non-Hispanic, (3) Black/African-American, non-Hispanic, (4) Hispanic, no race specified, (5) Hispanic, race specified, (6) More than one race, non-Hispanic, (7) Native Hawaiian/Pacific Islander, non-Hispanic, (8) White, non-

Hispanic, (-9) Missing. There were 1,006 missing data out of 23,503 students. Respondents who had missing values for race/ethnicity were removed from the study.

After selecting only students who enrolled into a bachelor's or associate's degree program in 2013, the racial groups analyzed in this study excluded America Indian/Alaska Native, non Hispanic and Native Hawaiian/Pacific Islander students due to a small number (78 students) with the acknowledgement of 1.1% America Indian/Alaska Native alone and 0.2% Native Hawaiian/Pacific Islander of all people living in the United States (Jones et al., 2021). The Hispanic students in group 4 and group 5 were merged to be a single group of Hispanics. I recoded the five racial categories to be analyzed: (0) White, (1) Asian, (2) Black, (3) Hispanic, (4) More than one race.

SES is a continuous composite variable within the range from (-1.75) to 2.57. This composite variable was generated through principal components factor analysis (weighted by W1STUDENT) and standardized to a mean of 0 and standard deviation of 1 on a z-scored scale. The SES variable was created based on five items of the highest education among parents/guardians in the two-parent family of a responding student, or the education of the sole parent/guardian (X1PAR1EDU), the education level of the other parent/guardian in the two-parent family (X1PAR2EDU), the highest occupation prestige score among parents/guardians in the two-parent family of a responding student or the prestige score of the sole parent/guardian (X1PAR1OCC2), the occupation prestige score of the other parent/guardian in the two-parent family (X1PAR2OCC2); and 5. family income (X1FAMINCOME) in 2008 (Ingels et al., 2011, p.168). The negative scores represent students whose families are lower in SES than the mean for the population (i.e., 2009 ninth graders), and the positive scores are those whose families are higher than mean SES for the population.

Based on the conceptual framework, high-school math and science attainment value, math and science utility value, math and science self-efficacy, and high school GPA are hypothetically strong predictors of STEM college major choice. Furthermore, adding college academic achievements including STEM credits earned and STEM GPA are predictors of STEM major choice decision-making. According to theories, environmental factors including collegiate supports and personal circumstances are related to the probability of STEM college major choice. These predictors are in below details.

Predictor variables

High school Math and Science attainment value, utility value, self-efficacy

The subjective task values and expectancy self-efficacy were measured on Likert-scales of 1 to 4 in which 1 was “Strongly agree” to 4 was “Strongly disagree” for six variables (math attainment value, math utility value, math self-efficacy, science attainment value, science utility value, and science self-efficacy). These variables’ composite scales were constructed by NCES from multiple survey items using principal components factor analysis (weighted by W1STUDENT) and standardized to a mean of 0 and standard deviation of 1. These composite continuous variables were reversely coded by NCES so that the higher scores represent a greater sense of attainment value, utility value, and self-efficacy.

Math attainment value. This scale is a composite of two items which asked students to what extent they saw themselves as a math person and to what extent other people saw them as a math person. NCES measured these scales to have Cronbach’s alphas of .84 (Ingels et al., 2011). Items were recoded by NCES as appropriate so that higher scores indicated a greater sense of identity.

*Math utility value***.** This scale is a composite of three items which asked students how much they agreed that math was useful for everyday life, college, or a future career. NCES measured these scales to have Cronbach's alphas of .78 (Ingels et al., 2011).

Math self-efficacy. This scale is a composite of four items which asked students how confident they were that they could do an excellent job on tests, understand the most difficult material presented in the textbook, master skills, and do an excellent job on assignments in their mathematics courses. NCES measured these scales to have Cronbach's alphas of .90 (Ingels et al., 2011).

Science attainment value. This scale is a composite of two items which asked students to what extent they saw themselves as a science person and to what extent other people saw them as a science person. NCES measured these scales to have Cronbach's alphas of .83 (Ingels et al., 2011).

Science utility value. This scale is a composite of three items which asked students how much they agreed that science was be useful for everyday life, college, or a future career. NCES measured these scales to have Cronbach's alphas of .75 (Ingels et al., 2011).

Science self-efficacy. This scale is a composite of four items which asked students how confident they were that they could do an excellent job on tests, understand the most difficult material presented in the textbook, master skills, and do an excellent job on assignments in their science courses. NCES measured these scales to have Cronbach's alphas of .88 (Ingels et al., 2011).

<u>High School GPA</u> is a continuous variable with the minimum score of .25 and maximum score of 4. The high school GPA reflects the actual grade point average of a student's high school grades, and not another measurement.

Collegiate factors

STEM credits earned is a continuous variable with the range of 0 – 115 credits using the NCES grant definition of science, technology, engineering, and mathematics. Dual enrollment is excluded from this variable. This variable utilizes "normalized" credits, which places the hours or credit units on a common scale so that credit units can be compared across institutions. Duplicate course records, created by the transfer of course credits to one or more additional institutions, are only counted once. Courses with a normalized grade of advanced placement, audit, or drop are excluded from this calculation (HSLS:09 Codebook).

STEM GPA variable indicates known grade-point average (GPA) for all STEM courses, using NCES grant definition of STEM, courses taken during the student's postsecondary education as of June 2016. This continuous variable ranges 0-4.

Besides the high school and college academic achievements, collegiate supports are hypothesized to predict whether students select a STEM major or not.

Collegiate support: Merit-aid received. The sum of only grants and scholarships students received when enrolled at the primary first year institution. This amount includes state merit-only grants and scholarships, and institutional merit-only grants and scholarships, including athletic scholarships. Among all 10,619 college enrolled students in either bachelor's or associate's program level in 2013, 45.2% respondents had missing values for this question. Of those with data on this item, only 10.6% students received merit aid in the range $5,992-38,607, while 89.4% did not receive merit aid.

Collegiate support: Need-aid received. Total amount of need-based grants received while students enrolled at the primary institution during the first academic year attended postsecondary education after high school. This amount includes the sum of federal Pell Grants, the Federal Supplemental

Educational Opportunity Grant awards, state need-based grants, and institutional need-based grants. The total needs aid received within 0 and $50,340. Among 10,619 students enrolled in bachelor's or associate's degree program, 47. 6% missing data and 26.4% students received need aid.

Collegiate support: Research with faculty. Students were asked about whether they participated in research project with a faculty member as a part of their college education by the end of February 2016. The binary variable was created by NCES, 1 indicated "Yes", and 0 indicated "No".

Collegiate support: Academic services used. Students were asked whether they used academic support service such as tutoring or writing centers by the end of February 2016 (i.e. students had visited, emailed, or in any way communicated with and received information or help from a school office or department that offers a particular service counts as use of that service). The binary variable was created by NCES, 1 indicated "Yes", and 0 indicated "No".

Collegiate support: Career planning or job placement services. Students were asked whether they used career planning or job placement services such as tutoring or writing centers by the end of February 2016 (i.e. students had visited, emailed, or in any way communicated with and received information or help from a school office or department that offers a particular service counts as use of that service). The binary variable was created by NCES, 1 indicated "Yes", and 0 indicated "No".

<u>Personal circumstances</u>

Collegiate supports are considered external factors in student decision to pursue a STEM major or not, besides high school and college learning experiences and achievements, their personal circumstances matter in their college major decisions. These circumstances are individual daily considerations and have effects on their college education career.

Work schedule interfered with academic performance during college. Students were asked their level of agreement that their work schedule interfered with their academic performance in college. This variable refers to who worked for pay while attending college sometime between July 2012 and February 2016. The Likert-scale 1-4 for (1) Strongly agree to (4) Strongly disagree. I recoded (1) for whom agreed with the statement and (0) for whom disagreed with to create a binary variable.

Money worry for regular expenses in 2015. Students were asked in calendar year 2015 when students were in the period of their second and third college year about whether they ever worried about having enough money for regular expenses. The binary variable was created by NCES, 1 indicated "Yes", and 0 indicated "No". The respondents reported whether "worry about having enough money for regular expenses?".

Disability. This variable indicates if any of the following had ever been true for the respondent, that meant they had disability or special need: 1) had a serious difficulty concentrating, remembering, or making decisions, 2) had been told by a health or education professional that he/she had ADHD or ADD (Attention Deficit Hyperactivity Disorder or Attention Deficit Disorder), 3) had a learning disability, 4) was deaf or had a serious difficulty hearing, 5) was blind or had a serious difficulty seeing, or 6) had any other disability or special need. The binary variable was created by NCES, 1 indicated "Yes", and 0 indicated "No".

<u>Analytic Sample</u>

The HSLS:09 public dataset was downloaded from the NCES website to identify the relevant variables of interest for the research questions under study. There were 23,503 students in the public dataset. I restricted the data to create an analytic sample based on number of students enrolled in college bachelor's or associate's degree programs in 2013 and their first major declared in their third college year 2016. Specifically, the analytical sample was processed as follows:

1) respondents enrolled in a bachelor's program or an associate's degree program in 2013;

Among 23,503 students in the public HSLS:09 dataset, there were 10,140 cases with missing data. 13,363 respondents had data in categories 1-7 as noted below and I selected those with responses of categories 1-3 in which categories 2 and 3 were grouped together as university transfer associate's degree. Seven categories NCES coded based on the first follow-up data collection wave in 2013 for this variable X3PROGLEVEL (1) Bachelor's degree program, (2) University transfer Associate's degree program, (3) Other Associate's degree program, (4) Certificate or diploma program (occupational training), (5) No program, just taking courses, (6) Other, (7) Don't know. There were 41 students responded in group 3 "Other Associate's degree program" that declared a STEM major in 2016 data collection wave, so I selected groups 1, 2, 3 for this study. I coded a binary variable with (0) = "bachelor's degree program", (1) = "university transfer associate's degree program". There were 10,619 eligible students retained.

2) racial observations with small number of observations and with missing data were dropped; 78 students from America Indians/Alaska Natives and Hawaiian Natives/Pacific Islanders and 378 missing data were dropped. There were 10,163 eligible students retained.

3) given the significant role of family SES, I dropped 354 missing data of SES. There were 9,809 eligible students retained.

4) respondents declared the first major field of study in 2016 defined by NCES. There were 2,319 students were coded missing data by NCES, so these students were removed, and 7,490 students retained.

After checking missing data of the entire sample, there were 32 observations with missing data in 10 or more variables of the analysis, so these observations were removed.

The final sample size was 7,458 students.

Weights

The sample data were weighted by multi-round postsecondary transcript data (W5W1W2W3W4PSTRANS), and postsecondary student records data (W5W1W2W3W4PSRECORDS) provided by NCES (Duprey et al., 2020, p.59).

Two classes of approaches to address analytical issues from complex samples are “design-based” for single-level, or student-level analyses, or “model-based approach for multilevel modeling. This study applied design-based approaches to get correct variance estimates and apply balanced repeated replication (BRR) (Rutkowski et al., 2010; Thomas & Heck, 2001). BRR variance estimation can be conducted with HSLS:09 public-use data (Ingels et al., 2011, p.172). The BRR weights are constructed to capture the variation associated with the sampling information. The BRR weights are of 200 weight variables that are used to correct the standard errors for the sampling plan to “produce estimates that are representative of the HSLS:09 target population of students for each study round and component” (Duprey et al., 2020, p. 48).

As HSLS:09 is a complex survey data, the “svyset” command in STATA indicates the survey data for estimating accurate population parameters and standard errors and I indicate svyset BRR at the beginning for the analysis involved either student postsecondary transcripts or student postsecondary records (Duprey et al., 2020).

HSLS:09 manuals guides “NCES standards require unit nonresponse bias analyses to be conducted when weighted unit response rates fall below 85 percent. The base weights account for differential selection probabilities. Analysis weights are constructed by adjusting the base weights for unknown eligibility, if applicable, and nonresponse to mitigate bias induced by those who did not respond to the study. The weights are further calibrated to known population totals to construct analysis weights which enable population estimates to be calculated from sample data." (Duprey et al., 2020, p.46). A set of 200 BRR weights used to analyze five data collection waves of postsecondary transcript data (W5W1W2W3W4PSTRANS001–200), and of postsecondary student records (W5W1W2W3W4PSRECORDS001–200) with financial aid information.

Depending on the available observed data, the population size varies. For example, the mean of high school GPA is 3.07 with standard errors of .02 for the sample of 10,366 students (who enrolled in bachelor's or associate's program) representing 2,098,533 population size. Similarly, the mean of college STEM GPA is 2.50 with standard errors of .02 for the sample of 10, 042 (who enrolled in bachelor's or associate's program) representing 1,964,034 population size.

<u>Analytic strategy</u>

To provide an overview of STEM major choice, I first conducted descriptive analyses. In addition, I used logistic regression to examine the extent to which predictors were associated with students' odds of declaring a STEM major as the first major field of study after controlling for gender, race, SES as the base model. Logistic regression models produce odds ratios for independent variables; these odds reflect the increase or decrease in the likelihood of the outcome (i.e., STEM or non-STEM major choice) for every one -unit increase in the independent variable (Trusty & Niles, 2003). I used the logit regression rather than the linear probability model (LPM) because LPM need to satisfy statistical assumptions of normality, linearity, and continuity for OLS regression which is difficult to achieve particularly when dependent variable is dichotomous (Mehmetoglu & Jakobsen, 2017). Thus, it is more appropriate to analyze dichotomous dependent variable through the logit transformation.

In this study, the following blocks of data were used to add variables for analysis:

Block 1 (2009): student gender, race, SES

Block2 (2009): adds math_attainment, math_utility, math_self-efficacy, science_attainment, science_utility, science_self-efficacy

Block 3 (2013): adds student high school GPA

Block 4 (2016): adds collegiate supports, personal circumstances

Block 5 (2016): adds college credits earned in STEM, college STEM GPA.

Analysis process

There are two sets of analysis. For the first research question, the first analysis with no missing data applied different "pweight" for (i) transcripts data, and (ii) student records data with financial aid information. The non-missing sample, 3,071 students, was analyzed with both "pweight". Although I intended to compare coefficients from analysis using listwise deletion and multiple imputation, I encountered a STATA software difficulty in that the application was unable to analyze complex survey data with BRR for imputed data. Thus, I did not proceed further with imputed data analysis as a means of addressing missing data. The second analysis using pairwise deletion for transcripts data only resulted in 4,350 students. This permitted examining any different pattern than the non-missing sample. For the second research question, pairwise deletion is used with Transcripts data only. The sample was 4,086 students. The third question has an analytical sample of 5,324 students.

Assumptions

For both research questions, assumptions of logistic regressions were tested for (i) linearity, (ii) absence of multicollinearity. For research question 3, regression assumptions of linear regressions include (i) normally distributed errors; (ii) normally distributed data, (iii) absence of multicollinearity. The assumptions were tested in Figure 10 and Table 20. Regression diagnostics for influential observations include (i) leverage, (ii) DFbeta, (iii) Cook's distance (Valliant et al., 2009). These options are disabled in STATA for complex survey data, so cannot detect influential observations. This will be examined in the future using R program instead with "*svydiags*" package.

Limitations

First, utilizing a secondary public data only allows to apply constructs and variables that were designed for the survey data collection rather than explaining other aspects the researcher wants to explore. For example, the binary variable "academic support services" is a single-indicator method to categorize whether students used or not, rather than the frequency of usage of services such as mentoring, writing center to understand specifically what STEM-domain-related activities, or how much efforts students have invested toward their STEM major decision. Moreover, students have different socialization experiences in each discipline, such as internships are more popularly required by engineering rather than math, so the collegiate supports from this public longitudinal dataset did not capture those learning differences in determination to a major. The public dataset also did not include socializers' roles at college level to understand their influence in students' STEM major choice. For example, students have college instructors of the same gender and race, particularly among minority students for vicarious learning outcomes (Minaya, 2018; Sheu et al., 2018). Faculty become role model for students in exploring academic disciplines, and research reported access to mentors and role models has a positive impact on women's professional development in STEM, especially when the mentor is a woman (Alfred et al., 2019). Faculty were found the most important support for community college students through academic validation experiences and engineering transfers' interpersonal validation experiences because students were transformed with higher self-efficacy in their learning abilities (Zhang & Ozuna, 2015).

Second, previous research suggests that transferred students from a two-year college is quite different from their peers who directly enter a four-year university from high school due to transfer function of a community college to a four-year institution in engineering program (Zhang & Ozuna, 2015), and the transferred students were more likely to be low-income or first-generation students (Wyner et al., 2016). The sample of this study includes students enrolled into a primary university-transferred associate's program then transferred to a four-year institution in reaching a STEM major. The extent of research participation with faculty as a collegiate support factor may be varied for two-year transferred enrollees versus four-year college enrollees because students mostly took introductory courses at community colleges, while students at four-year liberal art or research institutions may have undergraduate research opportunities with faculty. The present study cannot explain beyond this binary variable about when students participated in research with faculty, at the primary enrolled institution or transferred institution, and how various research experiences among disciplines to develop student knowledge and skills in pursuing a STEM major.

Third, financial aid variable explains the sum of merit-aid and the sum of need-aid including athletic scholarships students received in their primary first year institution. The variable limits to understand within the scope of the primary first-year institution, so did not calculate the aid if whether students received in case to transfer to another institutions or states during their college education, and if they did not have any athletic scholarships versus students who received athletic scholarships.

Finally, the variable "had work schedule interfered with academic performance" did not show type of work whether related to the field of study or how many hours weekly students work in conflict with their college schedule to deepen understanding student situation in preparation for their major and career preparation.

CHAPTER FOUR

RESULTS

This chapter explains descriptive statistics and the regression results based on non-missing data for the first question with two separate data, one is Transcripts data and one is Student Records data with financial aid. The second and third research questions show regression results of pairwise deletion data.

Research question 1

(1) To what extent do high school math and science motivation and self-efficacy, collegiate factors, and personal circumstances promote or hinder students' STEM major choice, controlling for student background characteristics?

TRANSCRIPTS DATA

Descriptive statistics (n=3,071)

Figure 4. Representativeness of the sample for the population

```
. svyset

      pweight: w5w1w2w3w4pstrans
          VCE: brr
          MSE: on
    brrweight: w5pstrans001 .. w5pstrans200
  Single unit: missing
     Strata 1: <one>
         SU 1: <observations>
        FPC 1: <zero>

. sum  w5w1w2w3w4pstrans

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
w5w1w2w3w~ns |      3,071    301.1687    378.5393          0   6290.424
```

The "pweight" of transcripts data is [w5w1w2w3w4pstrans] and the summary statistics shows that each person in the sample represents anywhere from 0 to 6,290 people in the population. The non-missing data (n=3,071) represents 924,889 students in the population. The Appendix C shows the comparisons between excluded sample (n=10,619) and analytic sample (n=7,458), and the Appendix D provides descriptive statistics of non-missing models and full analytic sample. The first set of data analysis of non-missing values (n =3,071) for two sources of Transcripts data, and Postsecondary data were analyzed.

Declaration of STEM major

Overall, 23% students declared a STEM major as the first major field, and among these students, 57% were male. Female students enrolled higher at both bachelor's (BA) and associate's (AA) degree program levels, but higher male proportions declared STEM majors in both levels. In this sample, 60% students enrolled BA level and 40% enrolled AA level. Among students enrolled the BA program level, 28% students declared a STEM major, and 56% of them are male. Among students enrolled in the AA program level, 17% students declared a STEM major, and 57% of them are male (*see* Table 2). The racial group of students declared a STEM major was White (60.8%), the second group was Hispanic (15.1%), Black students (10.7%), Asian (7.8%), more than one race (5.5%) (*see* Table 3).

Sociodemographic characteristics

Among 3,071 students, the percentage of gender is of 57% female. Proportion of male students declared STEM majors larger than female students, and students in STEM majors had higher SES than non-STEM students. This sample has 38% students of color including (a) Asian (5%), (b) Black (9%), (c) Hispanic (16%), (d) Multi racial group (8%), while the White group is dominant (62%) (*see* Table 3). The mean SES is 0.28 with the standard error of 0.04, and male students have higher SES than female students (.30 vs. .26).

Students who enrolled into AA level have mean SES lower than whom enrolled a BA level (0.002 vs .46). Within BA level, there is slight difference in mean SES between male and female students (.47 vs. 46), but there is a larger gap between men and women enrolled at the associate level (.06 for men vs. (-.04) for women). Within male students, there is a huge gap of SES between students enrolled a BA level and a AA level (.47 vs. .06). A similar pattern is within female students (46 vs. (-.04)).

Students who declared a STEM major have higher SES than non-STEM students, specifically, STEM-declared students' average SES scores .47 (SE=.07) versus non-STEM students' average SES scores. .38 (SE=.08) (*see* Table 4). Within female student group, students declared STEM majors have higher SES, on average .35 (SE=.09), than non-STEM students .24 (SE=.05). Similarly, mean of SES of STEM male students is .40 (SE=.08) compared with mean of SES of non-STEM male students .26 (SE=.05).

Table 2. Student demographics by STEM and non-STEM majors

Variables (n=3,071)	Weighted proportion/mean (SE)	
	STEM	Non-STEM
Male	13 % (.01)	30 % (.02)
Female	9 % (.01)	48 % (.02)
Race		
Asian	2 % (.0)	3 % (.01)
Black	3 % (.01)	7 % (.01)
Hispanic	3 % (.01)	12 % (.02)
More than one race	1 % (.04)	6 % (.01)
White (reference)	14 % (.01)	49 % (.02)
SES	.42 (.06)	.28 (.04)

Table 3. STEM major choice by student background characteristics

Student background	Within gender	
Race	Male (SE)	Female (SE)
Asian	57% (.12)	43% (.12)
Black	28% (.14)	72% (.14)
Hispanic	74% (.10)	26% (.10)
More than one race	23% (.15)	77% (.15)
White (reference)	63% (.05)	37% (.05)
SES	.47 (.07)	.38 (.08)

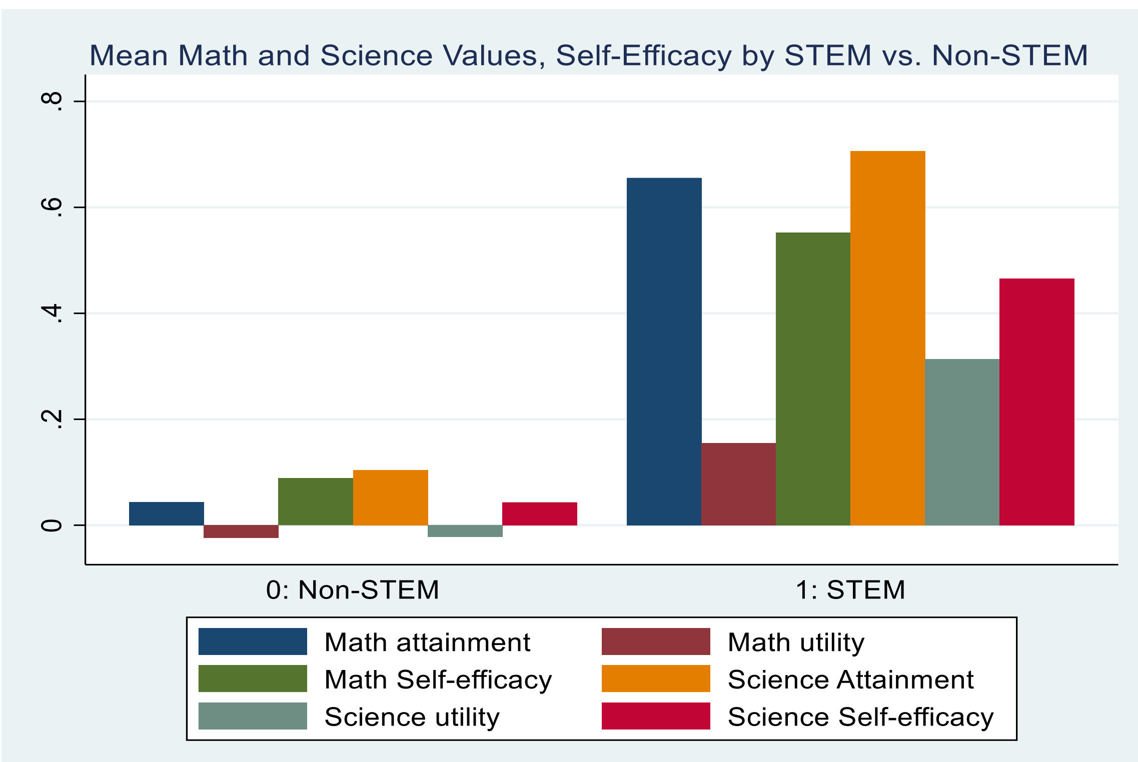
Math and Science attainment values, utility values, and self-efficacy between STEM vs. non-STEM students

As shown in Table 4, students who declared STEM majors have, on average, higher scores across all math attainment value (.71), math utility value (.16), math self-efficacy (.57) than non-STEM major declared students, .06, -.01, and .11 respectively. A similar pattern was found for science attainment (.72) value, science utility value (.29), and science self-efficacy (.48) of STEM majors declared students compared with non-STEM students, .10, -.03, and .06 respectively. Moreover, STEM-declared students have slightly higher HSGPA, earned triple the STEM credits in college, and higher STEM GPA than non-STEM students. Figure 5 illustrates the mean differences of math and science values and self-efficacy between STEM-declared and non-STEM declared students.

Table 4. Means of math and science attainment values, utility values, and self-efficacy, academic achievement variables by STEM and non-STEM majors

Variables (n=3,071)	Weighted mean (SE)	
	STEM	Non-STEM
Math attainment value	.71 (.08)	.06 (.06)
Math utility value	.16 (.06)	-.01 (.05)
Math self-efficacy	.57 (.06)	.11 (.04)
Science attainment value	.72 (.08)	.10 (.04)
Science utility value	.29 (.07)	-.03 (.06)
Science self-efficacy	.48 (.07)	.06 (.05)
High school GPA	3.32 (.06)	3.17 (.03)
College STEM credits earned	48.9 (2.23)	16.2 (.54)
College STEM GPA	2.75 (.07)	2.60 (.05)

Figure 5. Means of Math and Science attainment values, utility values, and self-efficacy by STEM vs. non-STEM



Math and Science attainment values, utility values, and self-efficacy, academic achievements between gender

In Table 5, compared math and science attainment values, math and science utility values, math and science self-efficacy, on average, male students have far higher math attainment value (.34), math self-efficacy (.40), science attainment value (.34), and science self-efficacy (.30) than female 9th graders with .11, .07, .17, and .05 respectively. Male students' average math utility values are also higher than female students, but equal scores in science utility value. Although male students have nearly five STEM credits earned more than female students, they have lower HSGPA and college STEM GPA than female peers. Female students have higher high school GPA than male students (3.30 vs. 3.07). When students enroll in college, male students, on average, earned more STEM credits than female students (26.32 vs. 21.41), but female students continue have higher mean college STEM GPA than male students (2.70 vs. 2.54).

Table 5. Means of math and science expectancy values and self-efficacy, academic achievement variables by gender.

Variables	Weighted mean (SE)	
	Male	Female
Math attainment value	.34 (.08)	.11 (.05)
Math utility value	.11 (.06)	-.03 (.05)
Math self-efficacy	.40 (.06)	.07 (.05)
Science attainment value	.34 (.07)	.17 (.05)
Science utility value	.04 (.09)	.04 (.05)
Science self-efficacy	.30 (.07)	.05 (.04)
HSGPA	3.07 (.05)	3.30 (.03)
College STEM credits earned	26.32 (1.33)	21.41 (1.08)
College STEM GPA	2.54 (.07)	2.70 (.06)

Math and Science attainment values, utility values, and self-efficacy between enrolled program level

Students enrolled the BA level have higher scores in math and science attainment value, self-efficacy, and science utility value than students enrolled in the AA level. Notably, students enrolled in the AA level have far higher math utility than students enrolled in the BA level (*see* Table 6). Students enrolled bachelor's program level have higher high school GPA (3.36 vs. 2.94), college STEM credits earned (26.69 vs. 18.15) and college STEM GPA (2.83 vs. 2.31) than students enrolled in the university transfer program.

Table 6. Means of math and science expectancy values and self-efficacy, academic achievement variables by enrolled degree program level BA vs. AA

Variables	Weighted mean (SE)	
	Bachelor's	Associate's
Math attainment	.29 (.05)	.07 (.08)
Math utility	-.05 (.04)	.16 (.07)
Math self-efficacy	.23 (.04)	.18 (.07)
Science attainment	.34 (.04)	.09 (.07)
Science utility	.07 (.05)	-.004 (.10)
Science self-efficacy	.26 (.04)	.04 (.07)
HSGPA	3.36 (.03)	2.94 (.06)
College STEM credits earned	26.69 (1.15)	18.15 (1.04)
College STEM GPA	2.83 (.04)	2.31 (.09)

College STEM credits earned and STEM GPA by race

Asian students have the highest mean of college STEM credits earned (35.42), while Black and students from more than one race have the lowest mean (18.20) (*see* Table 7). Black students have lowest STEM GPA among five racial groups at 1.96 scores, while other four groups' scores range 2.28-2.93 (*see* Table 8).

Table 7. Mean of college STEM credits earned at college across race

Variables	Weighted Mean	SE	[95% CI]
White	25.13	1.12	22.92-27.34
Asian	35.42	3.77	27.98-42.85
Black	18.20	2.40	13.46-22.94
Hispanic	19.03	2.14	14.81-23.25
More than one race	18.20	2.40	13.48-22.93

Table 8. Mean of college STEM GPA at college across race

Variables	Weighted Mean	SE	[95% CI]
White	2.80	.05	2.69-2.90
Asian	2.93	.11	2.72-3.14
Black	1.96	.20	1.56-2.35
Hispanic	2.28	.14	1.20-2.56
More than one race	2.64	.12	2.41-2.87

Collegiate Supports and Personal Circumstances

Nearly a fifth of students (19%) participated in research with faculty in college, more than half (54.5%) used academic services on campus, 34% used career services (*see* Appendix D). Among personal circumstances during college career, 44% agreed their work schedule interferes with their academic performance, notably 58% worried about money for regular expenses in 2015, and nearly a third (28%) of students reported having a disability.

In this sample, 54.4% students used academic services on campus. Among those used academic services, 25.7% declared STEM majors. Students selected STEM majors have higher collegiate supports of faculty research participation (31%), academic services (60%), and career services (39.4%) used than non-STEM students (*see* Table 9). STEM-declared students also have higher proportion in work schedule and academic performance interference (45.4%), but lower money worry about regular expenses in 2015 (51.6%), and disability than non-STEM students (21.6%).

Table 9. Cross tabulation of categorical variables of collegiate supports and personal circumstances by STEM

Variables	Code	Major	
		Non-STEM	STEM
Research with faculty	1	15.6%	31%
Academic services used	1	52.7%	60%
Career services used	1	32.2%	39.4%
Work schedule & academic performance interference	1	43.5%	45.4%
Money worry about regular expenses in 2015	1	59.5%	51.6%
Disability	1	30.5%	21.6%

Female students have higher proportions in five variables of collegiate supports and personal circumstances than male students, except the work schedule and academic performance interference (*see* Table 10). For example, among 3071 students, 20.3% female students participated in faculty research. More female students reported using academic services and career services (60.4% and 35.1%) than male students (46.5% and 32.2%). Furthermore, among personal circumstances, female students have lower agreement with work schedule and academic performance interference, but higher money worry about regular expenses in 2015, and disability than male students.

Table 10. Cross tabulation of categorical variables of collegiate supports and personal circumstances by Gender

Variables	Code	Gender	
		Male	Female
Research with faculty	1	17.7%	20.3%
Academic services used	1	46.5%	60.4%
Career services used	1	32.2%	35.1%
Work schedule & academic performance interference	1	44.4%	43.7%
Money worry about regular expenses in 2015	1	52.1%	61.8%
Disability	1	24.3%	31.5%

Logistic regression with non-missing data

In chapter three, I illustrated the model building process I used to estimate the probability of students declaring their first major in STEM or not. Results from the logistic regression assumptions indicated no multicollinearity and acceptable adherence to linearity. Table 11 shows the variance inflation factor (VIF) and the reported range of VIF is 1.04-1.82 to detect inflated standard errors due to multicollinearity. As the threshold of VIF values are greater than 5 and 1/VIF are less than 0.2, the estimated coefficients become less stable (Mehmetoglu & Jakobsen, 2017), so none of the predictors are highly correlated, in other words, there is no problem with multicollinearity in this model. Furthermore, all interactions of nine continuous predictors and the log of itself have significant values greater than .05, indicating that the assumption of linearity of the logit have been met for six math and science values and expectancies, and three academic achievements of HSGPA, college STEM credits earned, and STEM GPA.

Table 11. Variance Inflation Factor Values for Predictors in Model

Variables	VIF	1/VIF
Female	1.17	0.86
Asia	1.04	0.96
Black	1.20	0.84
Hispanic	1.22	0.82
More than one race	1.08	0.93
SES	1.20	0.83
Math attainment value	1.69	0.59
Math utility value	1.34	0.74
Math self-efficacy	1.69	0.59
Science attainment value	1.55	0.65
Science utility value	1.50	0.66
Science self-efficacy	1.54	0.65
HSGPA	1.82	0.55
College STEM credits earned	1.37	0.73
College STEM GPA	1.63	0.61
Research with faculty	1.12	0.89
Academic services used	1.12	0.89
Career services used	1.10	0.91
Work schedule and academic performance interference	1.13	0.89
Money worry about regular expenses in 2015	1.12	0.90
Disability	1.12	0.89

The base line model (Model 0) includes student background characteristics of gender, race, and SES. Model 0_a tests the interactions of gender and race, gender and SES, race and SES, and found no evidence of interaction effects, so this model result is omitted in the table 12. Model 1 adds student 9th grade math attainment value, math utility value, math self-efficacy, science attainment value, science utility value, and science self-efficacy. Model 2 adds student HSGPA, then Model 3 adds collegiate supports, Model 4 adds personal circumstances, Model 5 adds college STEM credits, and Model 6 adds college STEM GPA. Table 12 shows the moderate positive correlations between HSGPA and college STEM credits (r=.35, p<.05), and between college

STEM credits earned and college STEM GPA (r=.30, p <.05), a strong positive correlation between HSGPA and college STEM GPA (r=.55, p<.05).

Table 12. Pairwise correlations of unweighted continuous variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) SES	1.00									
(2) Math attainment value	.11*	1.00								
(3) Math utility value	-.05*	.29*	1.00							
(4) Math self-efficacy	.09*	.57*	.34*	1.00						
(5) Science attainment value	.15*	.30*	.14*	.17*	1.00					
(6) Science utility value	.01	.17*	.41*	.16*	.42*	1.00				
(7) Science self-efficacy	.12*	.27*	.17*	.37*	.49*	.34*	1.00			
(8) HSGPA	.22*	.28*	-.02	.19*	.14*	.04*	.17*	1.00		
(9) College STEM credits earned	.17*	.30*	.09*	.22*	.25*	.14*	.21*	.35*	1.00	
(10) College STEM GPA	.19*	.19*	-.04*	.11*	.07*	-.02	.07*	.55*	.30*	1.00

**p <.05*

The "*nestreg*" command does Wald tests on each model (Acock, 2018; Williams, 2021). "nestreg" fits nested models by sequentially adding blocks of variables and then reports comparison tests between the nested models (Acock, 2014).

Table 13. Logistic regression results of six models

DV = STEM	Model1 (OR) (SE)	Model2 (OR) (SE)	Model3 (OR) (SE)	Model4 (OR) (SE)	Model5 (OR) (SE)	Model6 (OR) (SE)
Female	.50*** (.09)	.58*** (.10)	.51*** (.09)	.48*** (.08)	.50*** (.09)	.58*** (.12)
Asian	2.12** (.62)	1.98 (.75)	2.05 (.75)	2.04 (.78)	2.00 (.76)	1.29 (.38)
Black	1.57 (.49)	1.71 (.47)	2.17** (.64)	1.87** (.56)	1.85** (.56)	2.14** (.81)
Hispanic	1.05 (.36)	1.13 (.38)	1.26 (.43)	1.20 (.39)	1.18 (.40)	1.14 (.40)
More than one race	.76 (.23)	.63 (.22)	.67 (.23)	.65 (.22)	.67 (.23)	.91 (.36)
SES	1.26 (.16)	1.09 (.14)	1.04 (.13)	.99 (.13)	1.00 (.14)	1.02 (.18)
Math attainment		1.51*** (.16)	1.42** (.15)	1.44** (.15)	1.44** (.15)	1.37** (.19)
Math utility		.87 (.10)	.89 (.10)	.90 (.10)	.89 (.10)	.89 (.11)
Math self-efficacy		1.28** (.14)	1.29** (.14)	1.26** (.14)	1.26** (.15)	1.25 (.18)
Science attainment		1.62*** (.19)	1.64*** (.19)	1.63*** (.19)	1.66*** (.19)	1.56** (.21)
Science utility		1.09 (.13)	1.08 (.12)	1.048 (.11)	1.05 (.11)	.99 (.14)
Science self-efficacy		1.02 (.11)	1.00 (.11)	1.02 (.11)	1.01 (.11)	.90 (.11)
HSGPA			1.51** (.27)	1.38 (.24)	1.33 (.25)	.82 (.20)
Research with faculty				1.95*** (.36)	1.93*** (.38)	1.06 (.25)
Academic services used				1.48** (.28)	1.54** (.29)	1.43 (.32)
Career services used				1.00 (.19)	1.00 (.19)	.95 (.23)
Work schedule & academic performance interference					1.38 (.25)	1.50 (.33)
Money worry for regular expenses					.82 (12)	.93 (19)

Table 13 Continued

Disability					.69	.72
					(.14)	(.18)
College STEM credits earned						1.08***
						(.01)
College STEM GPA						.62***
						(.08)
Constant	.38***	.24***	.07***	.07***	.07**	.07**
	(.05)	(.05)	(.04)	(.05)		
n	3,071	3,071	3,071	3,071	3,071	3,071
F	(6,194) =5.11	(12,188) = 9.93	(13,187) = 8.65	(16,184) =8.2	(18,181) =7.83	(21,179) =8.24
p-value	<.001	<.001	<.001	<.001	<.001	<.001

*** *p<.01, ** p<.05, * p<.1*

Overall, the Hosmer and Lemeshow test for the final model (Block 6) yielded F(9,191)=1.11, p >.05 indicating no evidence of lack of fit, or suggesting the model is good fit for these data (Archer & Lemeshow, 2006). Model 6 has F(21,179), p<.001 suggesting that the main effects model is a significant improvement over the model 0 including only gender, race, SES. Based on this model fit, female, Black, math attainment value, science attainment value, college STEM credits earned, and college STEM GPA are positively significant predictors of the probability of STEM major choice, controlling for other variables constant.

Notably, without STEM college credits earned and STEM GPA in model 5, female (compared to male), Black (compared to White), math attainment value, math self-efficacy, research with faculty, academic services used on campus, academic used on campus are predictive of the probability of STEM major selection. Specifically, females are predicted 50% less likely to declare STEM majors than males holding all other independent variables constant [95% CI: .39 - .86]. All else equal, a Black student is expected to have 1.85 greater odds of declaring a STEM major than the odds of a White student in the third-year college. Particularly, students participated with faculty research are expected a 1.93 increase in the odds of STEM major selection, and students who used academic services on campus are expected a 1.54 increase in the odds of STEM major selection.

However, when college STEM credits earned and college STEM GPA are added to model 6, math self-efficacy, research with faculty, and academic services used on campus variables are not statistically significant. With one unit increase in college STEM credits earned, the odds of STEM credits earned in college for students declared STEM majors are estimated to be .08 times the odds of STEM credits earned in college for non-STEM students [95% CI: 1.06-1.09], holding all other independent variables constant. Holding all else constant, for every one unit increase in college STEM GPA, the predicted odds for STEM majors decreased by 38% [95% CI: .49-.80]. For every one-unit increase in math attainment value, the odds of STEM major declaration increased about 37% than non-STEM majors students, holding all other independent variables constant [95% CI: 1.03-1.81]. Similarly with science attainment value, for every one unit increase in science attainment, the odds of selecting a STEM major increased 56% [95% CI: 1.20 to 2.02], all else equal.

STUDENT RECORDS DATA

The key difference in this model and those previously presented is the inclusion of student financial aid measures, namely whether the student received need and/or merit aid at their first primary institution. This examination required using a different dataset as these variables are not included in the transcript data but only in the student records dataset, available publicly from the National Center for Education Statistics.

The pweight is w5w1w2w3w4psrecords results in summary statistics shows that each person in the sample represents anywhere from 0 to 9,879 people in the population.

Descriptive statistics

There are slight differences in proportions of categorical variables between Student Records data and Transcripts data as illustrated in the Table 14. Notably, the Student Records data provide information on financial aid that shows more students enrolled into a BA level in 2013 than an AA level in contrast with Transcripts data. Many continuous variables have higher means and SE in Student Records data than in Transcripts data, except science utility variable in Student Records data is .02 lower score than in Transcripts data.

Table 14. Mean and Proportion of Categorical and Continuous variables

	STUDENT RECORDS DATA	TRANSCRIPTS DATA
Categorical variables		
Enrolled		
Bachelor’s program level	63%	60%
Associate’s program level	37%	40%
Non-STEM (1)	78%	77%
STEM (1)	22%	23%
Male (0)	43%	43%
Female (1)	57%	57%
Asian	4.5%	5%
Black	9.4%	9%
Hispanic	15.1%	16%
More than one race	7.4%	8%
White	64%	62%
Research with faculty		
(0)	81%	81%
(1)	19%	19%
Academic services used		
(0)	44%	46%
(1)	56%	54%
Career services used		
(0)	65%	66%
(1)	35%	34%
Money worry about regular expenses in 2015		
(0)	43%	42%
(1)	57%	58%

Table 14 (continued)

Disability			
(0)	72%	72%	
(1)	28%	28%	
Continuous variables	M (SE)	M	SE
SES	.31 (.04)	.28	.04
Math attainment	.21 (.05)	.19	.04
Math utility	.03 (.04)	.02	.04
Math self-efficacy	.21 (.03)	.20	.03
Science attainment	.24 (.04)	.24	.03
Science utility	.04 (.05)	.06	.04
Science self-efficacy	.15 (.04)	.14	.03
HS_GPA	3.20 (.03)	3.18	.03
Merit aid	.29 (.02)		
Need aid	.46 (.03)		
College-STEM credits earned	23.50 (.88)	23.51	.86
College_STEM GPA	2.63 (.05)	2.60	.05

pweight = w5w1w2w3w4psrecords

There are 1,022 missing values of both merit aid and need aid. Another 79 observations have missing values of need aid variable. As financial aid data are collected from both institution level and student level. NCES coded (-8) used for all student-level variables when a sample member is a nonrespondent to either PETS or SR, (-6) used when the value reported by the institution was outside the valid range for that field, and (-9) used for questions that are not answered within a survey when the respondent was eligible for the question (Duprey et al., 2020, p.105). Merit aid variable has (-8) and (-6) missing values. Need aid variable has (-9), (-8), and (-6) missing values. Respondents with these missing values were removed from the regression analyses. The sample size was 1,970 students.

<u>Logistic regression results (n=1,970)</u>

The population size is 918,963 students. The interactions of gender-race, gender-SES, race-SES were tested and found not statistically significant, so model 1 including interaction terms is not displayed in Table 15. Logistic regression estimates result.

Building upon five blocks to examine whether factors of collegiate supports and personal circumstances influence the probability of STEM major choice after college STEM credits earned and college STEM GPA were added to the proceeding blocks of student background characteristics and other high-school level variables. The "*nestreg*" command does Wald tests on each model (Acock, 2018; Williams, 2021). "nestreg" fits nested models by sequentially adding blocks of variables and then reports comparison tests between the nested models (Acock, 2014).

Table 15. Logistic regressions of five models using Students Records data (n=1970)

DV = STEM	M1 (OR) (SE)	M2 (OR) (SE)	M3 (OR) (SE)	M4 (OR) (SE)	M5 (OR) (SE)	M5: 95% CI
Female	.45*	.52*	.46*	.60	.56*	.32 - .97
	(.08)	(.10)	(.10)	(.16)	(.16)	
Asian	2.04*	1.82	1.86	.79	.73	.32-1.67
	(.71)	(.72)	(.72)	(.30)	(.31)	
Black	1.63	1.76	2.29*	2.24	2.07	.86 – 4.97
	(.58)	(.60)	(.83)	(1.01)	(.92)	
Hispanic	.98	1.13	1.27	1.07	.98	.46 – 2.10
	(.33)	(.38)	(.42)	(.40)	(.38)	
More than one race	.75	.65	.68	.83	.86	.26 – 2.86
	(.27)	(.25)	(.26)	(.49)	(.52)	
SES	1.28	1.07	1.01	.90	.92	.64 – 1.32
	(.15)	(.13)	(.12)	(.15)	(.17)	
Math attainment		1.63*	1.52*	1.43*	1.44*	1.02-2.02
		(.23)	(.21)	(.25)	(.25)	
Math utility		.87	.88	.86	.85	.64 – 1.13
		(.11)	(.11)	(.12)	(.12)	
Math self-efficacy		1.18	1.20	1.24	1.21	.86 – 1.70
		(.15)	(.15)	(.21)	(.21)	

Table 15
(continued)

Science attainment		1.64* (.23)	1.68* (.24)	1.61* (.29)	1.63* (.29)	1.15 – 2.31
Science utility		1.06 (.13)	1.04 (.13)	.96 (.16)	.97 (.17)	.69 – 1.37
Science self-efficacy		1.06 (.14)	1.03 (.14)	.99 (.17)	1.00 (.17)	.72 -1.40
HSGPA			1.54* (.32)	.81 (.26)	.73 (.22)	.40 -1.31
College STEM credits earned				1.09* (.01)	1.09* (.01)	1.07 – 1.12
College STEM GPA				.53* (.09)	.55* (.10)	.38 - .78
Merit aid					1.04 (.25)	.65 – 1.66
Need aid					1.04 (.16)	.77 – 1.42
Research with faculty					1.01 (.29)	.58 -1.79
Academic services used					1.59 (.43)	.93 -2.71
Career services used					1.03 (.33)	.55 – 1.92
Work schedule & academic performance interference					1.60 (.42)	.95 - 2.67
Money worry for regular expenses in 2015					1.07 (.31)	.60 -1.90
Disability					.49 (.19)	.23 -1.04
Constant	.39 (.05)	.24 (.05)	.06 (.04)	.19 (.17)	.18 (.17)	
n	1970	1970	1970	1970	1970	
F	(6,194) =5.35	(12,188) = 7.69	(13,187) =6.79	(15,185) = 11.39	(23,177) =7.41	
p-value	<.001	<.001	<.001	<.001	<.001	

*** *p<.01*, ** *p<.05*, * *p<.1*

The F statistics for the first block, 5.35, is for a test of the joint significance of the first block of student background characteristics to explain variance in STEM major choice. The F statistics for the second block, 8.94, for a test of the joint significance of the second block variables (math and science values and expectancy) in a logistic regression of both the first and second blocks of variables. Similarly, F-statistics for the third (4.38), fourth (41.52), and fifth (1.28) block respectively added HSGPA, collegiate STEM credits earned and college STEM GPA, collegiate supports and personal circumstances. Each block explained a statistically significant increase in the variance of STEM major choice.

The F-statistics for model specification whether the observed 0/1 values on the outcome variable match the expected 0/1 values, $F(9,191) = 1.16$, $p > .05$ indicates there is evidence of an adequately-fit model (M5).

Gender, math attainment, science attainment, college STEM credits earned, and college STEM GPA were statistically significant predictors of the outcome variable STEM/non-STEM. There was statistically significant difference between female and male students, specifically, being females compared with males decreases the odds of STEM declared majors by a factor of .56 holding all other variables constant (95% CI: .32 - .97). The odds ratio for math attainment 1.44 means the odds of STEM majors declared increase by a factor of 1.44 with every unit increase in math attainment, in other words, the higher math attainment, the greater the likelihood students' declared STEM a major, holding all other variables constant (95% CI: 1.02-2.02). The predicted odds of STEM major declaration increase by 63% per unit increase in science attainment, holding all other variables constant (95% CI: 1.15 – 2.31). The predicted odds of STEM major declaration increase by 9% per unit of STEM credit students earned in college, holding all other variables constant (95% CI: 1.07 – 1.12). For every unit increase in college STEM GPA, the predicted odds of declaring a STEM major (STEM=1) decreased by 47%.

In sum, compared with non-missing TRANSCRIPTS data, this non-missing STUDENTRECORDS data, neither merit nor need based aid were associated with STEM major choice.

Across these two datasets, gender, math attainment value, science attainment value, college STEM credits earned, college STEM GPA are predictors of STEM major choice. Females in these datasets are less likely to have declared STEM majors, even though they have higher STEM GPA than male peers. Notably, when college STEM GPA increased, the expected odds of STEM majors declared decreased between 38-47%. The more STEM credits students earned in college, the higher odds of students declaring a STEM major, ranging from 8-9% with each unit increase in STEM credits earned across the two data. Interestingly, the higher math attainment and science attainment values, the higher probability of students choosing a STEM major.

Using "*margins*" command, the predicted probability of choosing a STEM major is higher for males than females (.25 vs. .20). The predicted probability between gender and choosing a STEM major varies by students' race/ethnicity. Specifically, the probability of a hypothetical Black male student choosing a STEM major is 33% whereas that probability is 26% for Black female students. Similarly for other racial groups of students, Asian male students have a 22% chance, while Asian female students would have a 17% chance. Hispanic male students have a 25% chance compared with 20% chance of Hispanic female students. For more than once race group, male students have a 23% chance, whereas female students have a 19% chance. Finally, White male students have a predicted probability of 25%, while White female students have a 20%

chance. Graphs below shows predict margins of gender by math attainment, science attainment, and the different probability between gender.

Figures 6. Predictive margins of gender by five predictors

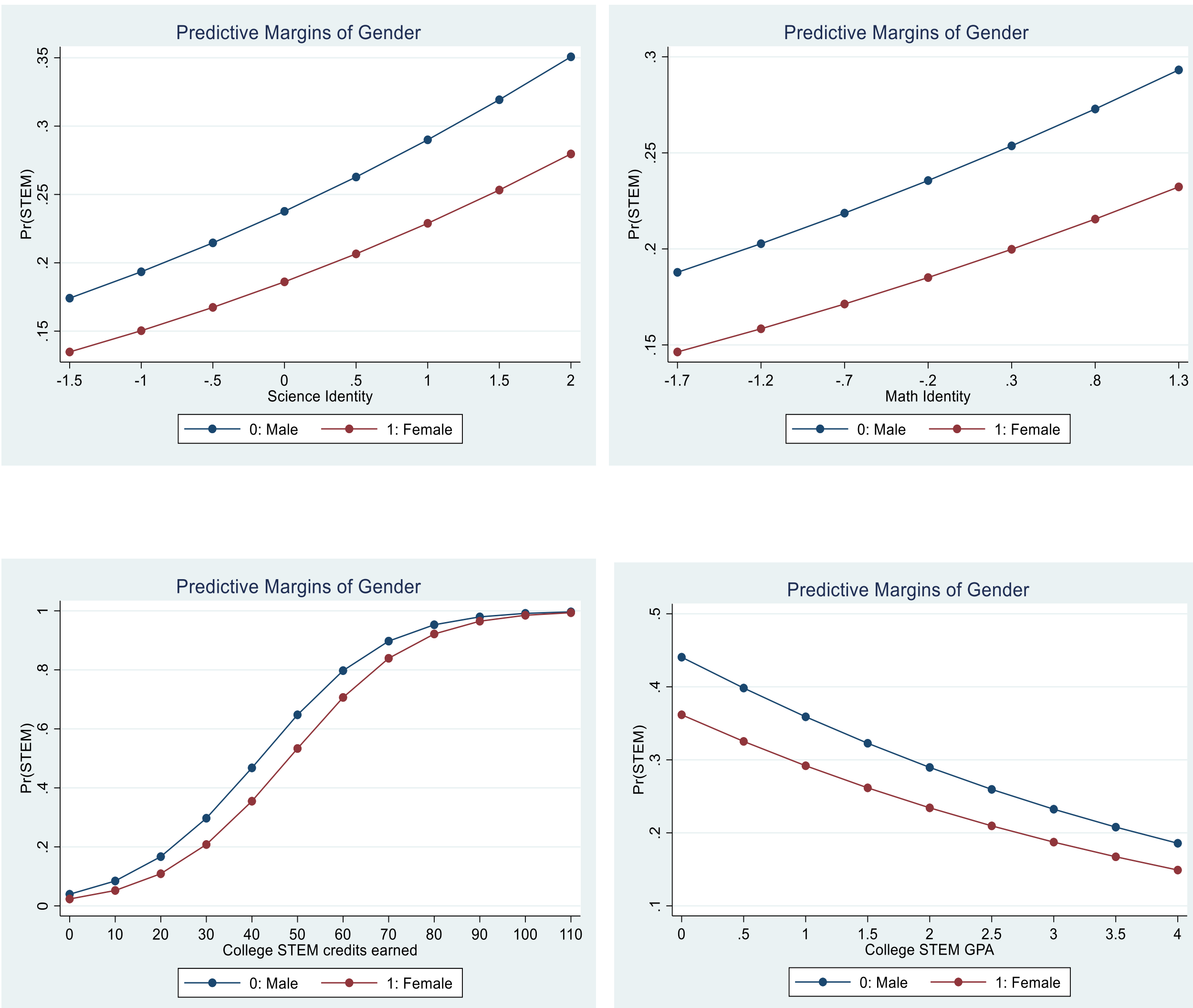
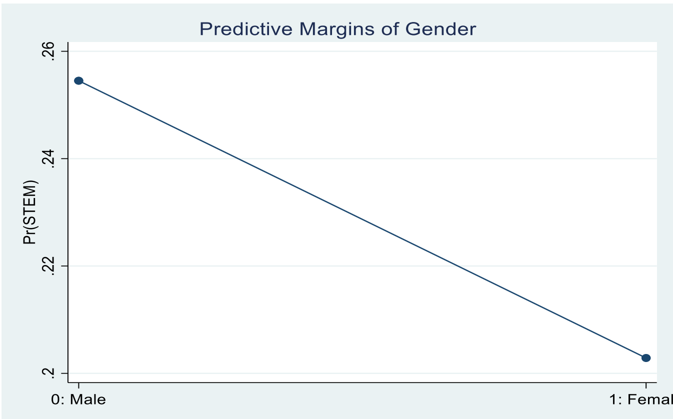


Figure 6 Continued



SES, math utility, math self-efficacy, science utility, and science self-efficacy were found not statistically significant in predicting the probability of STEM major declaration. Furthermore, none of contextual factors including collegiate supports and personal circumstances were statistically significant. In all, math and science attainment values, gender, and college academic achievements are predictors of STEM major choice.

Logistic regression with pairwise deletion (n=4,350)

As I explained in chapter three that STATA program currently does not support to variance estimate with *public* complex survey data with BRR weights for multiple imputation (i.e., mi estimate: svy: logistic (vce = brr)), I ran the last logistic regression of pairwise deletion among 7,458 students enrolled both BA's and AA's program levels for the first research question. Table 18 shows the result of model 5 only, including

Table 16. Logistic regression

STEM	Odds Ratio	SE	[95% CI	]	Sig
Female	.58	.12	.39	.86	***
Asian	1.29	.38	.72	2.32	
Black	2.14	.81	1.02	4.51	**
Hispanic	1.14	.40	.57	2.28	
More than 1 race	.91	.36	.42	1.98	
SES	1.02	.18	.72	1.44	
Math attainment	1.37	.20	1.03	1.81	**
Math utility	.89	.11	.70	1.13	
Math self-efficacy	1.25	.18	.93	1.67	
Science attainment	1.56	.21	1.20	2.03	***
Science utility	.99	.14	.75	1.30	
Science self-efficacy	.90	.11	.70	1.16	
HSGPA	.82	.20	.51	1.32	
College STEM credits earned	1.08	.01	1.06	1.09	***
College STEM GPA	.62	.08	.48	.80	***
Research with faculty	1.06	.25	.67	1.69	
Academic services used	1.43	.32	.91	2.23	
Career services used	.95	.23	.60	1.53	
Work schedule and academic performance interference	1.50	.33	.97	2.31	*
Money worry for regular expenses in 2015	.93	.19	.62	1.40	
Disability	.72	.18	.45	1.16	
Constant	.17	.14	.03	.86	**

*** *p<.01*, ** *p<.05*, * *p<.1*

The model test F(9, 191) = 1.11, p > .05 indicating the observed data fit with the model. Similar with the complete non-missing data (n=3,071), there is evidence that gender, math attainment value, science attainment value, college STEM credits earned, STEM GPA are predictors for the probability of STEM major selection. The predicted odds of STEM major declaration increase by 8% per unit of STEM credit students earned in college, holding all other variables constant (95% CI: 1.06 – 1.09). For every unit increase in college STEM GPA, the predicted odds of declaring a STEM major (STEM=1) decreased by 38%, holding all other variables constant (95% CI: .48-.80).

Research question 2

To what extent do collegiate factors and personal circumstances predict the probability of STEM major choice, controlling for student background characteristics?

The first research question answered the probability estimates of STEM major declaration based on both high-school and college variables, while the second research question seeks the answer for the probability of STEM major declaration based on college-level variables only. In the next section, the logistic regression results are presented.

Logistic regression

As the SCCT choice model explains the more confident students are in a specific domain, the higher likelihood they will choose that domain for their career. As postsecondary education is an important transition for students to highly skilled workforce, particularly in STEM fields of which require more than 70% at least bachelor's degrees for future occupations, the present study considered STEM credits earned as a proxy for students' confidence and mastery in specific discipline. Research question 2 emphasizes the college period and examines how STEM credits earned, STEM GPA, collegiate supports, and personal circumstances relate to the likelihood of STEM major choice students declared in their third-year college. Based on the model building approach described in Chapter three, the logistic regression analysis addresses this research question. The sample was 4,086 out of 7,485 observations. The F-test (9, 191) = 1.18, $p>.05$ indicating the observed data fit the model.

Table 17 shows gender, college STEM credits earned, and college STEM GPA are predictors for the likelihood of STEM major choice. The third research question examines further what variables affect STEM GPA that is proxy of student academic success. The odds ratio of gender less than 1 indicating female students are less likely to select STEM majors than male students. The predicted odds of STEM major declaration increase by 8% per unit of STEM credit students earned in college, holding all other variables constant (95% CI: 1.07 – 1.10). For every unit increase in college STEM GPA, the predicted odds of declaring a STEM major (STEM=1) decreased by 40%, holding all other variables constant (95% CI: .49 – .73).

Table 17. College-level logistic regressions (n=4,086)

DV = STEM	Odds ratio	SE	95% CI
Female	.50***	.08	.36 - .69
Asian	1.17	.39	.61 - 2.26
Black	1.52	.50	.79 – 2.92
Hispanic	.96	.25	.57 – 1.62
More than one race	1.02	.32	.56 – 1.88
SES	1.02	.15	.77 -1.35
Research with faculty	1.14	.26	.73 – 1.78
Academic services used	1.33	.23	.95 – 1.88
Career services used	.98	.21	.64 – 1.49
Work schedule & academic performance interference	1.45	.28	.99 – 2.12
Money worry for regular expenses in 2015	.91	.15	.66 -1.26
Disability	.79	.15	.54 – 1.16
College STEM credits earned	1.08***	.01	1.07 – 1.10
College STEM GPA	.60***	.06	.49 -.73
Constant	.13***	.05	.06 - .26
n	4,086		
Table 18 Continued			
F	(14,186) = 17.31		
p-value	<.001		

*** *p<.01*, ** *p<.05*, * *p<.1*

Research question 3

What factors predict college STEM GPA?

The findings in research questions 1 and 2 indicate the negative association of college STEM GPA in estimating the probability of STEM major declaration. The third research question investigates further what variables affect college STEM GPA. As college STEM GPA is a continuous variable, ranging from 0 to 4, the following analytical approach is linear regression with assumptions tested for linearity and no multicollinearity.

<u>Linear regression</u>

The multiple regression in Table 18 shows three models including student demographic characteristics (model 1), adding collegiate supports and personal circumstances (model 2), and STEM credits earned (model 3) for predicting college STEM GPA. Model 3 has $F(13, 187) = 23.71$, $p<.05$, $R^2 =.24$, the model explains 24% of the variance in college STEM GPA that significant increases 6% from model 2 without STEM credits earned $F(12, 188) = 24.74$, $p<.05$, $R^2=.18$. The sample of model 3, with pairwise deletion, includes 5,324 students who represent a population size of 1,243,199.

Gender, Black, Hispanic, more than one race, SES, research with faculty, career services used, work schedule interferes with academic performance, disability, and college STEM credits earned were statistically significant. Figures 7-9 illustrate the predictive margins of college STEM GPA by gender, race, research with faculty, career services used, work schedule interferes with academic performance, and disability. Black, Hispanic, and more than one race were negatively associated with STEM GPA. Average STEM GPA increases .17, .20 points and decreases .25 and 16 points respectively for every unit increase in participating faculty research, using career services on campus, work schedule and academic performance interference, and disability. The interactions between gender and race, gender and SES, race and SES were not statistically significant, that means college STEM GPA scores by race do not vary by gender, SES.

The effect sizes of Black is negatively strong (b=-.69, SE=.12, 95% CI [(-.92) – (-.46)]), Hispanic is negative moderate (b=-.30, SE=12, 95% CI [(-.53)-(-.07)]), and more than one race is negative weak (b=-.19, SE=.00, 95% CI[(-.38) – (-.001)]), holding all else constant. Research with faculty has positive weak relationship with the average mean of STEM GPA (b=.17, SE=.05, 95% CI [.08 - .26)]), holding all else constant. Career services used on campus is positive moderate relationship with STEM GPA (b=.20, SE=.05, 95% CI [.10 - .30)]), holding all else constant. When students reported their work schedule interfered with their academic performance, the mean STEM GPA decreases .25 points (b=-.25, SE=.06, 95% CI [(-.37 – (-.14)]), holding all else constant. Students who reported having a disability had lower mean STEM GPA by .16 points (b=-.16, SE=.05, 95% CI [(-.27 – (-.06)]), holding all else constant.

Table 18. Linear regressions for college STEM GPA

	Model 1	Model 2	Model3
Female	0.19^{***}	0.20^{**}	0.24^{***}
	(0.05)	(0.06)	(0.06)
Asian	0.11	0.11	0.01
	(0.06)	(0.08)	(0.06)
Black	-0.74^{***}	-0.77^{***}	-0.69^{***}
	(0.10)	(0.12)	(0.12)
Hispanic	-0.32^{**}	-0.34^{**}	-0.30^{*}
	(0.12)	(0.13)	(0.12)
More than 1 race	-0.34^{**}	-0.25^{*}	-0.19^{*}
	(0.10)	(0.10)	(0.09)
SES	0.21^{***}	0.12^{**}	0.09^{*}
	(0.04)	(0.04)	(0.04)
Research with faculty		0.32^{***}	0.17^{***}
		(0.05)	(0.05)
Academic services used		-0.04	-0.06
		(0.05)	(0.05)

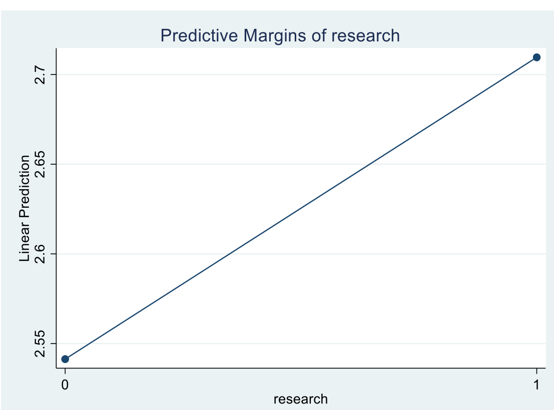
Table 18 continued

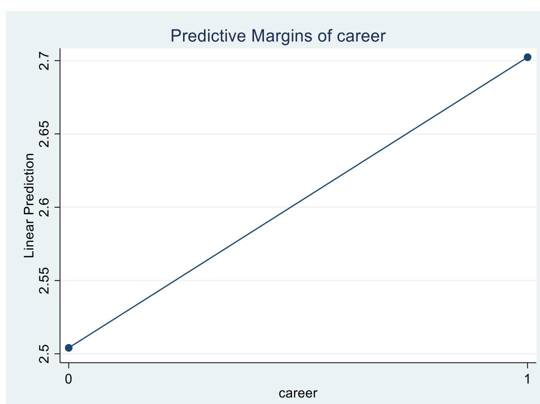
Career service used		0.22***	0.20***
		(0.05)	(0.05)
Work schedule and academic performance interference		-0.27***	-0.25***
		(0.06)	(0.06)
Money worry in 2015		-0.12*	-0.08
		(0.05)	(0.05)
Disability		-0.21***	-0.16**
		(0.05)	(0.05)
College STEM credits earned			0.01***
			(0.00)
_cons	2.55***	2.71***	2.41***
	(0.04)	(0.07)	(0.08)
N	7010	5324	5324
F	(6, 194) = 26.38	(12, 188) = 24.74	(13, 187) = 28.71
P_value	*p<.05*	*p<.05*	*p<.05*
R^2	0.11	0.18	0.24

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The below predictive margins graphs illustrate the different effects of gender, race, collegiate supports, and personal circumstances in relation to the probability of STEM major choice.

Figures 7. Predictive Margins of College STEM GPA by Collegiate Supports





Figures *8*. Predictive Margins of College STEM GPA by Personal Circumstances

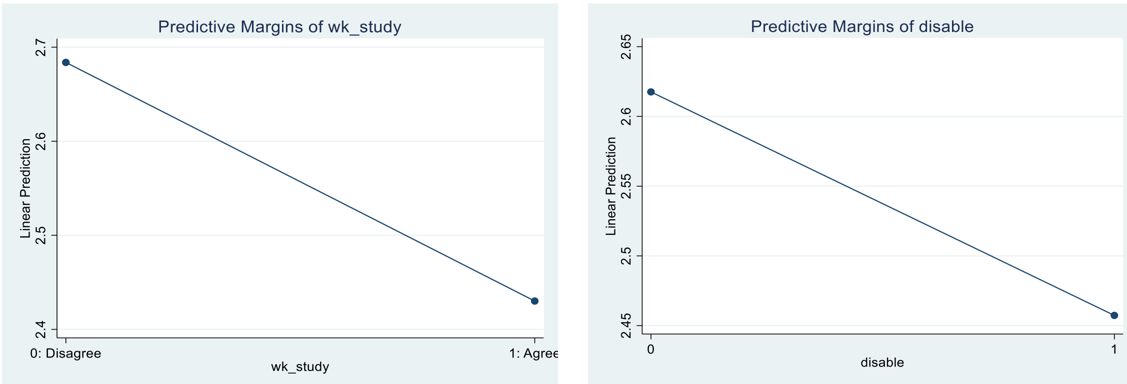


Figure 10 indicates the assumptions are met due to the normal distributed errors of College STEM GPA as an outcome variable, the normal distribution data (Q-Q plot), no multicollinearity. The VIF values do not exceed 5 and none of 1/VIF is lower than .20. (*see* Table 19)

Figure 9. Linear regression assumptions

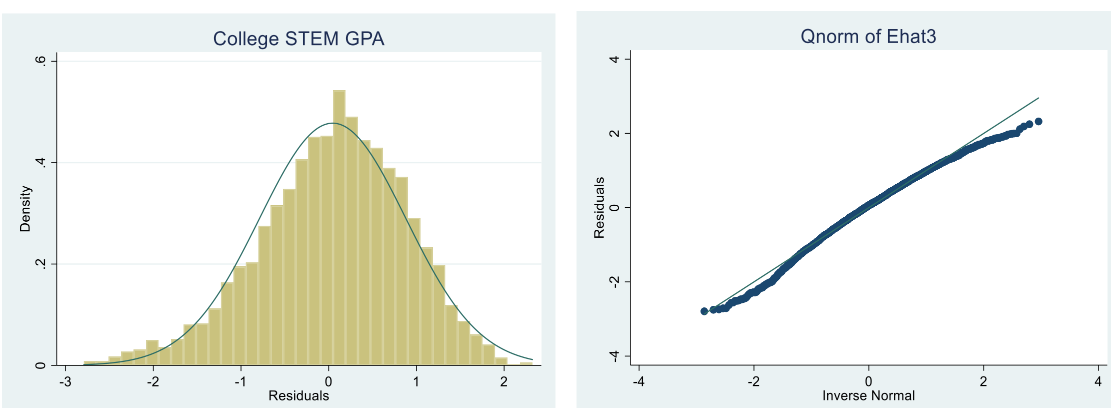


Table 19. Variance Inflation Factor (VIF) Values for Predictors in Model

	VIF	1/VIF
SES	1.13	0.88
College STEM credits earned	1.13	0.89
Money worry for regular expenses in 2015	1.11	0.90
Work schedule and academic performance interference	1.11	0.91
Academic services used	1.10	0.91
Race	1.09	0.92
Research with faculty	1.08	0.93
Career services used	1.07	0.93
Disability	1.07	0.93
Gender	1.04	0.96

This chapter provides series of data analysis to answer three research questions. For the first research questions, two sets of non-missing data using Transcripts data (n=3,071) and Student Records data with financial aid (n=1,970) were used. Results indicated gender, math attainment value, science attainment value, college STEM credits earned, and college STEM GPA predicted the probability of choosing a STEM major. The higher math and science attainments, the greater the likelihood students selected STEM majors. There was no statistically significant evidence that merit or need aid had any relationship with STEM major choice.

A further analysis of pairwise logistic regressions using Transcripts data (n=4,350) reports the similar patterns with non-missing dataset. For the second research question examining college-period-variables, gender and college STEM credits earned, STEM GPA are predictors of STEM majors. There was evidence that gender, race, SES, participation in faculty research, career services used in campus and STEM credits earned were positively associated with average STEM GPA, while work schedule and academic performance interference, and disability were negatively associated with the predicted mean of STEM GPA. Regression assumptions and postestimation estimates were examined. In the following chapter, these results will be discussed further.

CHAPTER FIVE

DISCUSSIONS, IMPLICATIONS, AND CONCLUSION

Introductions

STEM workforce demand is in more urgent need than ever globally. The US government and higher education institutions have collaborated over the last few decades to increase the supply side, as well as student college major choices aim to prepare for on-the-market and career. Historically, males have been dominant in STEM fields, while the near future demography forecasts the increasing current minority population in the US. Thus, supporting and fostering women and underrepresented minority students to pursue STEM fields would escalate the workforce supply as majority of STEM occupations require at least a bachelor's degree (Fayer et al., 2017; US Bureau of Labor Statistics, 2020).

Previous empirical studies have investigated high school-level factors such as math and science motivational beliefs (Jiang et al., 2020), math and science self-efficacy (Jewett, 2019; Kurban & Cabrera, 2020a) to predict student intention to major in STEM, or STEM major in third-year college. Few studies have examined both high school and college factors in relation to STEM college major choice. The present study employed two theories, SEVT and SCCT, to examine theoretically the association between 9th graders' math and science attainment values, math and science utility values, and math and science self-efficacy, high school and college academic achievements, collegiate supports, and personal circumstances during college career, and their STEM major choice in third-year college, accounting for student background characteristics. The data analysis used the most recent nationally representative dataset HSLS:09 to answer three research questions.

Discussions

Research question 1:

> To what extent do high school math and science motivation and self-efficacy, collegiate factors, and personal circumstances promote or hinder students' STEM major choice, controlling for student background characteristics?

Two publicly available datasets were analyzed for this question using BRR weights of the NCES public complex survey data as explained in chapter four.

The descriptive statistics reported greater female enrollment at both bachelor's (BA) and associate's (AA) degree program levels, but higher male proportions declared STEM majors in both levels. Students who declared STEM majors have higher SES than non-STEM students and male students who declared STEM majors have higher average SES than female students. Student enrolled in BA level have higher SES than those enrolled in AA level. More Black female students and female students of more than one race selected STEM majors than male students which contrasts with other racial patterns for Asian, Hispanic, and White students. More female students participated in research with faculty, used academic and career services on campus than male students. They also have higher proportions in personal circumstances of money worry about regular expenses in 2015 and disability than male students, but lower proportion of work schedule and academic performance interference than male students.

STEM-declared students have higher average scores in all six math and science variables (i.e. math attainment, math utility, math self-efficacy science attainment, science utility, and science self-efficacy) in comparison with non-STEM students. Students enrolled BA level have higher math and science attainment values, math and science utility values, math and science self-efficacy, and academic achievements in high school GPA, college STEM GPA, and college STEM credits earned than students enrolled in AA level. In a specific US rural context, Tran et al. (2021) reported students enrolled at a 4-year institution are more likely to consider STEM majors than students at a 2-year institution. Asian students, on average, earned the highest STEM credits and STEM GPA, while Black students have the lowest of college STEM credits earned and STEM GPA among five racial groups. Male students have, on average, higher math and science attainment values, math utility values, math and science self-efficacy, but the same science utility as female students from the non-missing data. Rangel et al. (2020) found evidence that female students are less positive about math than male counterparts, but they are more likely to be positive about science than about math. This study did not find that pattern, but found that females have higher high school GPA and college STEM GPA than male students while they earned less college STEM credits than male peers.

Among collegiate support factors, students who declared a STEM major have higher proportion in faculty research participation, academic services and career services used than non-STEM students (*see* Table 9). Comparing between gender of STEM majors – declared students of the non-missing dataset, a larger proportion of female students participated in faculty research and used academic services, career services on campus, disability than male STEM students (*see* Table 20). Among personal circumstances, when looking at gender and STEM majors, male students who major in STEM report greater work schedule and academic performance interference, money worry than female students majoring in STEM.

Table 20. Proportions of collegiate supports and personal circumstances of STEM students between gender (non-missing values)

STEM=1 (n=3,071)	Male (SE)	Female (SE)
Research with faculty	46% (.06)	54% (.06)
Academic services used	49% (.05)	51% (.05)
Career services used	47% (.05)	53% (.05)
Work schedule and academic performance interference	57% (.06)	43% (.06)
Money worry for regular expenses in 2015	53% (.05)	47% (.05)
Disability	44% (.07)	56% (.07)

Without college STEM credits earned and college STEM GPA, logistic regression results of non-missing data show gender, Black, math attainment value, math self-efficacy, science attainment value, research with faculty, and academic services used on campus were positive significant predictors of the STEM declared group. When college STEM credits earned and college STEM GPA were included in the logistic regression, gender, Black, math and science attainment, college STEM credits earned and college STEM GPA were predictors of the probability of STEM major group, math self-efficacy was not statistically significant. The higher math and science attainment values, the more likelihood students selected STEM majors. The odds of Black students (compared to White students) choosing a STEM major are 114% greater. With each credit increased in college STEM credits earned, the odds of STEM majors group was estimated to be .08 times than the odds of non-STEM students, holding all else constant. For every one unit increased in college STEM GPA, the predicted odds for STEM majors decreased by 38%. SES, college supports, and personal circumstances were not statistically significant in this analysis. Financial aid was not statistically significant. Inference statistics in the pairwise deletion logistic regressions illustrated the similar significant predictors of the probability of STEM major selection.

Female students in this sample have the same pattern with previous study in that they are less likely to declare a STEM major than male peers (Jewett, 2019, Moakler & Kim, 2014; Jiang et al., 2020; Kurban & Cabrera, 2020; Mau & Li, 2018; Nix & Perez-Felkner, 2020). The higher math ability, the more likely college students enrolled in STEM major (Wang & Degol, 2017). The most interesting finding of this study is the higher college STEM GPA students gained in the first three years of college education, the lower odds of STEM group students declared. This national representative sample shares a similar pattern of females' STEM GPA higher than males as of North Carolina state dataset (Stearns et al., 2020). In terms of racial comparison, like previous studies' findings of difference in STEM major choice between racial groups (Li, 2019; Nix & Perez-Felkner, 2019; Wolniak, 2016; Moakler and Kim, 2014), Black students had higher probabilities to declare a STEM majors than White peers, holding all else variables constant.

SES, math and science utility value, math and science self-efficacy were not predictors of the first major in STEM fields in this study, although self-efficacy was the predictor of academic satisfaction and intended persistence across the third and fourth semesters in engineering field (Fouad & Santana, 2017; Lent et al., 2015). Math and science - utility value, interest, and attainment value were significantly STEM career plans of White students (Gottlieb, 2018) but fewer significant relationships between these values and STEM career plans for Black and Hispanic students. Interestingly, the non-missing descriptive statistics reveal same average science utility between gender, and students enrolled AA's level had higher average math utility value than students enrolled BA's level. Furthermore, as Wang and her colleagues suggested that (Wang et al., 2019) successful transfer in STEM majors does not depend only with course-takings within

STEM subjects, but through credits students earn in coursework beyond STEM. I have tested model with college STEM credits attempted and college GPA instead of college STEM credits earned and STEM GPA, the findings were not consistent with Wang and her colleagues because there was evidence to reject the null hypothesis. Notably, math and science attainment values are strong predictors of STEM major selection over the period of seven years.

Based on the data analysis of non-missing data with and without financial aid, collegiate supports and personal circumstances were not statistically significant. These findings are similar with insignificant barrier factor of weekly work hours, but in contrast with financial aid support factor using ELS:2002 (Wang, 2013).

Research question 2:

> To what extent do collegiate factors and personal circumstances predict the probability of STEM major choice, controlling for student background characteristics?

The second research question examines college-level factors only that permits the researcher to see how college-level coefficients differ when high school factors were included and did not include in the regression models. Comparing odds ratio between model 6 in Table 13 and model 17, the odds ratio of gender without high-school factors was lower than the odds of gender with high-school factors, meaning the probability of STEM major choice was lower in only college-level model. Both college STEM credits earned, and college STEM GPA found statistically significant in both with and without-high-school-level models. Specifically, the odds ratio of college STEM credits earned were the same between two models (OR = 1.08). The odds ratio of college STEM GPA with high-school factors was slightly higher than the odds ratio of college STEM GPA without high-school level model (.62 vs. .60), but both were less than 1. In other words, the decreased likelihood of STEM major choice in only college-level was higher than in combined high school and college-level model. Notably, math and science attainment values were statistically significant in research question 1, showing that these variables could not be ignored to estimate the probability of STEM college major selection because of the positive association between the math and science attainment values, and STEM major declaration. In other words, the stronger math and science identity students perceived themselves as well as were perceived by others, the higher likelihood students pursued STEM college majors. The odds ratio of math attainment value was 1.37 and the odds of science attainment value was 1.56. For each unit increase of math attainment value, the odds of STEM majors declared increased 37%, and for each unit increase of science attainment value, the odds of STEM majors declared increased 5%. It is important for teachers, counsellors, parents, and other socializers to support high school students or earlier-stage students develop their math and science identities.

When female high school students have higher math and science identities that would enhance them to cope with the advanced STEM college courses. One of the reasons for STEM gender disparity found gender gaps in self-efficacy (Cheryan at al., 2017). Lehman et al. (2016) pointed out that freshman female students rated themselves lower in math ability, intellectual self-confidence than men in computer science, while female students took more college advanced courses in all major categories except STEM they took fewer (Shewach et al., 2019). Nix and Perez-Felkner (2019) analyzed ELS:2002 for math difficulty orientation and found the higher difficulty orientation, the higher probability of STEM majors college students selected, while Riegle-Crumb et al. reported high school senior male and female students were equally likely to take calculus course (Riegle-Crumb et al., 2012).

Thompson (2020) found math self-efficacy beliefs were not positively associated with high school calculus enrollment for White, Black students, but a negative association for Hispanic male students. Thompson also reported the math utility value was associated with calculus enrollment for only Hispanic females, and math attainment value predicted calculus enrollment for Black males. Jiang et al. (2020) found math subjective task value did not significantly predict STEM college majors, but science subjective task value was positively associated with STEM major choice. Although math and science utility values, and math and science self-efficacy were statistically insignificant in research question 1, high-school level math and science attainment values found strong predictors of STEM major declaration. That would elaborate student motivational beliefs and competence development in high school to grow and foster further demonstrated by college STEM credits earned and STEM GPA, together with gender, to be predictors of STEM major declaration in only college-level model of research question 2.

Female students in this study were less likely to declare STEM majors than male students, despite higher college enrollment rate, high school GPA, and STEM GPA. Moreover, the descriptive statistics showed STEM female students had higher rate of faculty research participation, academic and career services used on campus, and higher rate of disability than STEM male students (*see* Table 21). It is critical to understand high school math and science attainment values in the emerging adulthood of female students, in order to increase their self-awareness of competence and confidence to rate higher math and science scores. Also, it is worthwhile to explore whether the higher academic and career services on campus female students used would enable them to decide STEM majors based on math and science identities. If so, academic advisors and career counsellors on campus would support and provide students information beyond academic track and diverse STEM-related careers to continue growing their identities, and later, pursuing STEM careers, and reduce the changing major likelihood due to the lack of STEM career information offered (Marade & Brinthaupt, 2018).

The proportions of White, Asian, and Hispanic STEM male students were greater than female students, but more than double proportions of Black and more than one race female STEM students compared with male Black and more than one race STEM students (*see* Table 22). Given this context, Black and more than one race female students are expected to graduate in STEM majors, and later, to have STEM occupations in the future study when the next HSLS:09 data collection wave in 2025 will take place. In the next analytical discussion of research question 3, I examined what demographic characteristics and college-level factors affect student college STEM GPA.

Table 21. Proportions of STEM-declared students by gender

STEM=1 (n=3,071)	Male (SE)	Female (SE)
White	64% (.04)	36% (.04)
Asian	55% (.13)	45% (.13)
Black	29% (.11)	71% (.11)
Hispanic	57% (.12)	43% (.12)
More than one race	31% (.12)	69% (.12)

Research question 3:

What factors predict college STEM GPA?

The linear regression (*see* Table 19) showed gender, race except Asian, SES, research with faculty, career services used on campus, work schedule and academic performance interference, and disability were predictors of STEM GPA. Model 3 indicated that more than 24% of the variance in STEM GPA scores were predicted from those predictors, or a moderate model fit. Those predictors were consistent with previous empirical findings to understand relationships of student background characteristics, supports on campus, and personal conditions in relation to STEM major choice. A study of a large public university students in the US showed reasons why students changed a major were due to lack of sense of belonging in the academic major, or lack of sense of achievements such as negative grades, or a lack of knowledge about the specific field and the careers it offered (Marade & Brinthaupt, 2018). The present study founds students, particularly female STEM students, participated in faculty research higher than male STEM students to develop their knowledge and skills in STEM fields. Combined both BA's and AA's enrollees, among STEM Black students participated in faculty research, female students accounted for 68% (SE=.24, 95% CI = .21-1.14), 81% for Hispanic STEM female students (SE=.28, 95% CI = .25 – 1.37), 51% for more than one race students (SE=.25, 95% CI = .02 – 1.01) in comparison with male same racial peers. Among female STEM students participated in faculty research, the racial proportions were 62.1% White, 6.2% Asian, 10.2% Black, 16.3% Hispanic, and 5.2% more than once race. This result show underrepresented students had lower faculty research participation. The interactions between faculty and students varies by student background characteristics as Black students interacted more frequently with faculty for course-related matters than other racial groups, and they were less likely to assist faculty with research and most likely to report experiencing racial/ethnic discrimination (Park et al., 2020). Previous study findings that Black female engineering students experienced a more negative climate than their non-Black female peers (Rincon & George-Jackson, 2016), while engineering faculty played an important role for successfully transferred students from community colleges to a four-year institution because they encouraged and supported students scholar abilities and professional development (Zhang & Ozuna, 2015). The college student population has witnessed the increase of minorities to two-year colleges and many recent state policies expressly support transfer students from two-year colleges to four-year institutions (National Academies of Sciences, Engineering, and Medicine, 2016). Notably, Black students were more likely to participate in research when they attended institutions that offered undergraduate research experiences to first-year students as part of a structured program, and they were more likely to participate in undergraduate research programs than White students (Figueroa et al., 2013). Park et al.'s study (2020) reported female, Black, and Latinx students were more likely to leave STEM by the fourth year of college than male, White, and Asian American peers. As the present study sample declared their major in their third-year college career, future study examines the retention and graduation rates of female and underrepresented students to accomplish a STEM bachelor's degree.

Female STEM student used more career services on campus, while underrepresented students (i.e. Black, Hispanic, and more than one race) had negative relationships with the predicted STEM GPA. What math and science challenges in college students from Black, Hispanic, and more than one race groups face with to support them increase STEM GPA is an open question for further study. The coefficient of college STEM credits earned was positive significant 0.01 showing for each credit increase in STEM credits, only 1% increase in the STEM GPA. When students were in the first-year college, the higher STEM-related grades they gained, the higher probability they declared in biology and/or physical science/engineering major (Stearns et al., 2020). Thus, this study aggregated three-year STEM credits earned either at BA's program level or a university transfer's program level, future study examines to what extent STEM credits earned differ the likelihood of STEM major selection between students enrolled in BA's and AA's level.

SES, work schedule and academic performance interference, and disability were statistically significant in predicting college STEM GPA. For one unit SES increase, STEM GPA increased 9%. Previous studies showed low-income students were less likely to pursue STEM majors (Niu, 2017) but more likely to enroll at four-/two-year for-profit, or community colleges, while middle-/high-income counterparts were more likely to attend four-year institutions and selective colleges (Dynarski et al., 2018; Fry & Cilluffo, 2019; Oseguera & Hwang, 2014). This study descriptive statistics showed STEM students had higher SES than non-STEM students, and enrolled in BA's level higher than in AA's level (71% vs. 29%). Research on working during college have provided contradictory findings whether working during college were earnings for college education, food and housing expenses, and students who experienced working during college had higher earnings after graduation (Douglas & Attewell, 2019), but working during college cause lower academic achievements, retention and graduation rates, and longer time-to-degree because they had to split study and work time (Darolia, 2014). Chang et al. (2014) analyzed repeated-measures freshman and senior student surveys and found working full-time while attending college was negatively associated with underrepresented students' chances of persisting in a STEM field (Chang et al., 2014). In the present study, students who had work schedule interfered with academic performance were 25% lower STEM GPA than students wo did not have interference, particularly more male STEM students had the interference than female STEM students.

Using HSLS:09 dataset, Bittinger studied observations with disabilities in predicting their STEM major choice and reported female students had lower odds ratio than males in terms of the likelihood of pursuing a STEM major (Bittinger, 2018). The present study found students with disability had 16% lower STEM GPA than students without disability, and female STEM students had higher disability rate than male STEM students, but disability variable was not a predictor of the probability of STEM major choice.

The following sections will discuss implications for theories, practice, and future research.

Theoretical implications

The present study employs SEVT and SCCT to understand student motivations and behaviors in college major choice, and later, career preparedness. Both SEVT and SCCT have been used to study high school students (Jiang, et al., 2020; Marsh, 2020) and college students

(Alshahrani et al., 2018; Byars-Winston & Rogers, 2019). Out of four expectancy values of SEVT including math and science attainment value, math and science utility value that were examined in this study, only math and science attainment values supports SEVT to predict STEM major choice. Furthermore, in SCCT, the choice model explains the higher self-efficacy students gain in particular domains, specifically hereinafter math and science self-efficacy, they are more likely to choose STEM major (Sheu et al., 2018). Math and science self-efficacy constructs measure student excellence in subject tests, understanding of difficult problems in textbooks, mastery skills, and excellence in assignments. Together with math and science attainment values, only math self-efficacy was found statistically significant to predict STEM major choice in model 5 (*see* Table 13) when STEM credits earned and STEM GPA were not included in the regression. When STEM credits earned and STEM GPA added to the regression, math self-efficacy was not statistically significant that may explain their confidence in math strongly expressed by college-level academic achievements. In all, based on these theories, data in this study illustrate that students had high attainment values and confidence in high school math and science, the higher likelihood they took more college STEM credits, obtained higher college STEM GPA, and positively associated to their STEM major choice. The math and science utility values, math and science self-efficacy were found not to be predictors of the probability of STEM majors. As utility values and self-efficacy were measured in freshman high school of which the development stage of personal identity may be stronger at later stage of high school and college education, these values and expectancy may be deep-rooted into math and science attainment values as they grew, especially when students already enrolled college.

There was wide gap between math and science attainment values between gender, race, and college enrolled program levels, thus it is important to reduce these gaps in order to increase the diversity of future STEM workforce. Wigfield and Eccles (2020) explained “attainment value derives from the fit of perceived task characteristics with the individual’s core self-schema, social and personal identities, and ought selves, that is, the extent to which tasks allow or not the person to the manifest those behaviors that they view as central to their own sense of themselves, or allow them to express or confirm important aspects of self” (p.167). The earlier students perceive or expose themselves and are perceived by others to be good in math and sciences, the more likelihood they select STEM college majors, for example through the early summer STEM program, particularly for underrepresented students (Ashley et al., 2017; Kitchen et al., 2018; Oseguera et al., 2020), but these programs should be conducted at lower high school grades, rather than officially admitted stage and prior to the first college semester.

Implications for policy, practice, and future study

This study aims to understand what factors support or hinder the increase of women and underrepresented minority students in the future STEM workforce by investigating both high school and college factors, particularly among the third year first-major declared students. The study used the most recent national representative sample of the 9th graders in 2009 to analyze for the population size. Female students were less likely to declare STEM majors than male students as found in previous studies, even though they enrolled at rates higher than males in both bachelor’s or university-transfer program level, and have higher HSGPA and STEM GPA. However, female students earned less STEM credits than male peers, raising the question: does that limit female students to explore more in-depth STEM courses and gain confidence in the content? Moreover, female students had average high school math attainment value equals a third of the average male math attainment value, and female students’ science attainment value equals half of the average male science attainment value. Given the attainment and value measures are self-responded survey data, it would be interesting for future study to examine why girls rate their attainment values in math and science lower than boys, and how to increase their attainment values to be confident in these domains because both math and science attainment values were found predictors of college STEM major choice.

Female students participated in faculty research and used more career services on campus than male students. Research with faculty, and career services used were found to be predictive of STEM GPA that imply female encouragement and engagement to STEM fields in order to prepare for their future career. Institutions may disseminate more career services to female students, particularly underrepresented students because more female students from Black and more than one race racial groups in this study declared STEM majors than male peers. Providing exposure to a wide array of STEM jobs through career centers would increase female students' awareness and exploration to diverse STEM-related jobs, especially innovation leadership to engage higher career ladder and increase economic mobility as Wu et al. (2021) reported the higher innovation cultures in female CTOs' firms versus male CTOs. Moreover, supporting undergraduate female students to participate in faculty research progressively engages them to advance STEM knowledge and skills, though faculty in STEM fields at present are still male dominant. This learning experience would enhance their confidence and develop sense of belonging to STEM fields which may promote their pursuit of STEM majors and careers.

In terms of barriers, more female students reported having a disability in this study than male students, and higher STEM females with disability than male STEM students. It is critical to promulgate policies and interventions to support students with disabilities, particularly for female students with a disability who may experience "pre-set" double disadvantages to embark on future STEM workforce. Although merit aid and need aid were not statistically significant in this study, students with disability needed more financial supports for their learning tools and self-support equipment (Perlow et al., 2021). As the Obama era professional judgement guidance was rescinded in June 2020, nearly 40% of the website financial aid information did not reference disability in any way among 51 public four-year institutions observed, and fewer than one in five institutions provided clear steps for students to engage this policy. Furthermore, only 20% of institutional websites provided information of time limits to degree completion as federal aid allows 150% of the time allotted for degree completion. There is a need at federal-, state-, and institutional-level to support financially more grants or scholarships for female students with disability to reduce work-schedule as STEM female students had the work-study interference accounted for 43% compared with 57% STEM female students without disability who also had work-study interference. Thus, with nonloan aid, the higher likelihood students maintain high academic performance, and pursue STEM majors to prepare for future advancement of STEM professionals because the effects of nonloan aid is significantly associated with higher rates of college graduation and gives students flexibility to focus on studies and less time to work (Stoddard et al., 2018), especially a last three-decade successful example of Meyerhoff program at University of Maryland, Baltimore County.

Based on the current findings, rather than using STEM umbrella to describe various disciplines, it is worthwhile to investigate deeper how specific disciplines in STEM fields including science, computer science, engineering, math students have similar or different college experiences such as research with faculty, and financial aid supports to pursue their major. Also, future study explores what high school and college factors differ student likelihood of STEM major choice between students enrolled in bachelor’s program level and students enrolled in university transfer program level.

In terms of methods, extended quantitative methodological approaches using machine learning would be explored. Also, due to the current limited STATA programming options such as to estimate logistic regression for imputed data with BRR in HSLS:09, or to conduct regression diagnostics for survey data, this study can be extended further in the future upon the available STATA adds-on. Finally, the next wave of HSLS:09 scheduled to collect in 2025 when 9^{th} graders in 2009 reach the age of 30, the future research may provide additional information of student success based on graduation rate, or whether their occupations relate to college majors.

In conclusion, female and underrepresented students are competent to study STEM majors. School counsellors, teachers, academic advisors, career counsellors need to support female students to enhance, foster, and bold female's math and science attainment values in high school, so they have stronger beliefs to pursue STEM careers that is dominantly male at present. The future US innovations and competitiveness sustain when more females and underrepresented students engage in STEM fields.

REFERENCES CITED

Acock, A. C. (2014). STATA Manuals 13. In *A Gentle Introduction to STATA* (4th ed.). STATA Press. https://www.stata.com/manuals13/rnestreg.pdf

Acock, A. C. (2018). *A Gentle Introduction to STATA* (6th ed.). STATA Press.

Alfred, M. V., Ray, S. M., & Johnson, M. A. (2019). Advancing Women of Color in STEM: An Imperative for U.S. Global Competitiveness. *Advances in Developing Human Resources*, *21*(1), 114–132. https://doi.org/10.1177/1523422318814551

Alshahrani, A., Ross, I., & Wood, M. I. (2018). *Using social cognitive career theory to understand why students choose to study computer science*. 205–214. https://psycnet.apa.org/record/2018-49065-001

Alvarado, S. E., & Muniz, P. (2018). Racial and ethnic heterogeneity in the effect of MESA on AP STEM coursework and college STEM major aspirations. *Research in Higher Education*, *59*(7), 933–957.

Apriceno, M., Levy, S. R., & London, B. (2020). Mentorship during college transition predicts academic self-fffficacy and sense of belonging among STEM students. *Journal of College Student Development*, *61*(5), 643–648.

Archer, K. J., & Lemeshow, S. (2006). Goodness-of-fit Test for a Logistic Regression Model Fitted using Survey Sample Data. *The Stata Journal*, *6*(1), 97–105. https://doi.org/10.1177/1536867X0600600106

Arcidiacono, P., Aucejo, E. M., & Hotz, V. J. (2013). University differences in the graduation of minorities in STEM fields: Evidence from California. *American Economic Review*, *106*(3), 525–562.

Ashley, M., Cooper, K. M., Cala, J. M., & Brownell, S. E. (2017). Building better bridges into STEM: A synthesis of 25 years of literature on STEM summer bridge programs. *CBE—Life Sciences Education*, *16*(4), es3.

Biglan, A. (1973). Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology*, *57*(3), 204.

Bittinger, J. (2018). *STEM pipeline for students with disabilities: From high school to intentions to major in STEM*. https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=2293&context=dissertations_2

Blackburn, H. (2017). The Status of Women in STEM in Higher Education: A Review of the Literature 2007–2017. *Science & Technology Libraries*, *36*(3), 235–273. https://doi.org/10.1080/0194262X.2017.1371658

Bolds, T. (2017). *A Structural and Intersectional Analysis of High School Students' STEM Career Development Using a Social Cognitive Career Theory Framework*

Bottia, M. C., Mickelson, R. A., Jamil, C., Moniz, K., & Barry, L. (2021). Factors Associated With College STEM Participation of Racially Minoritized Students: A Synthesis of Research. *Review of Educational Research*, *91*(4), 614–648. https://doi.org/10.3102/00346543211012751

Briggs, C. (2014). *Mathematics: Self-efficacy, identity, and achievement among African American males from the high school longitudinal study*. Alabama State University.

Bureau of Labor Statistics. (2021). *Computer and Information Technology Occupations: Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics*. https://www.bls.gov/ooh/computer-and-information-technology/home.htm

Byars-Winston, A., & Rogers, J. G. (2019). Testing intersectionality of race/ethnicity× gender in a social–cognitive career theory model with science identity. *Journal of Counseling Psychology*, *66*(1), 30.

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

Cass, C. A., Hazari, Z., Cribbs, J., Sadler, P. M., & Sonnert, G. (2011). *Examining the impact of mathematics identity on the choice of engineering careers for male and female students*. F2H-1-F2H-5.

Castellanos, M. (2018). Examining Latinas' STEM Career Decision-Making Process: A Psychosociocultural Approach. *The Journal of Higher Education*, *89*(4), 527–552.

Catalyst. (2021). *Women of Color in the United States*. Catalyst. https://www.catalyst.org/research/women-of-color-in-the-united-states/

Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching*, *51*(5), 555–580. https://doi.org/10.1002/tea.21146

Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., & Bearman, S. (2011). The Role of Efficacy and Identity in Science Career Commitment Among Underrepresented Minority Students. *Journal of Social Issues*, *67*(3), 469–491. https://doi.org/10.1111/j.1540-4560.2011.01710.x

Chen, X. (2013). STEM Attrition: College Students' Paths into and out of STEM Fields. Statistical Analysis Report. NCES 2014-001. *National Center for Education Statistics*.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*(1), 1–35. https://doi.org/10.1037/bul0000052

Colby, S. L., & Ortman, J. M. (2015). *Projections of the Size and Composition of the U.S: 2014-2060*. US Census Bureau. https://www.census.gov/library/publications/2015/demo/p25-1143.html

Darolia, R. (2014). Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students | Elsevier Enhanced Reader. *Economics of Education Review*, *38*(2018), 38–50. https://doi.org/10.1016/j.econedurev.2013.10.004

Department of Education. (2017). *Data Point: Beginning College Students Who Change Their Majors Within 3 Years of Enrollment* (No. 2018–434). NCES. https://nces.ed.gov/pubs2018/2018434.pdf

Department of Education, N. (2019). *Digest of Education Statistics—STEM degrees conferred 2008-2018*. National Center for Education Statisics (NCES), Department of Education. https://nces.ed.gov/programs/digest/d19/tables/dt19_318.45.asp

Douglas, D., & Attewell, P. (2019). *Assessing the Impact of Student Work During College* (p. 51). Rutgers University. https://smlr.rutgers.edu/sites/smlr/files/Documents/Centers/EERC/eerc_working_students_012119.pdf

Douglas, D., & Salzman, H. (2019). Math Counts: Major and Gender Differences in College Mathematics Coursework. *The Journal of Higher Education*, 1–29.

Duprey, M. A., Pratt, D. J., Jewell, D. M., Cominole, M. B., Fritch, L. B., Ritchie, E. A., Rogers, J. E., Wescott, J. D., & Wilson, D. H. (2018). *High School Longitudinal Study of 2009 (HSLS:09) Base-Year to Second Follow-Up Data File Documentation*. NCES. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018140

Duprey, M. A., Pratt, D. J., Wilson, D. H., Jewell, D. M., Brown, D. S., Caves, L. R., Kinney, S. K., Mattox, T. L., Ritchie, N. S., & Rogers, J. E. (2020). High School Longitudinal Study of 2009 (HSLS: 09) Postsecondary Education Transcript Study and Student Financial Aid Records Collection. Data File Documentation. Appendices. NCES 2020-004. *National Center for Education Statistics*. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020004

Dynarski, S., Libassi, C. J., Michelmore, K., & Owen, S. (2018). *Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students* (No. w25349; p. w25349). National Bureau of Economic Research. https://doi.org/10.3386/w25349

Eccles, J. P., Adler, T., Futterman, R., Goff, S., Kaczala, C., Meece, J., & Midgley, C. (1983). Expectancies, values, and academic behaviors. *Achievement and Achievement Motives*, 75–146.

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Engberg, M. E., & Wolniak, G. (2013). *Teachers College Record: College Student Pathways to the STEM Disciplines*. *115*(010304). https://www.tcrecord.org/LIBRARY/exec.asp?ContentID=16740

Erwin, C., Binder, M., Miller, C., & Krause, K. (2021). Performance-Based Aid, Enhanced Advising, and the Income Gap in College Graduation: Evidence From a Randomized Controlled Trial. *Educational Evaluation and Policy Analysis*, *43*(1), 134–153. https://doi.org/10.3102/0162373720979180

Evans, C. A., Chen, R., & Hudes, R. P. (2020). Understanding Determinants for STEM Major Choice Among Students Beginning Community College. *Community College Review*, *48*(3), 227–251. https://doi.org/10.1177/0091552120917214

Fayer, S., Lacey, A., & Watson, A. (2017). *STEM occupations: Past, present, and future*. U.S. Bureau of Labor Statistics. https://www.bls.gov/spotlight/2017/science-technology-engineering-and-mathematics-stem-occupations-past-present-and-future/home.htm

Figueroa, T., Hughes, B., & Hurtado, S. (2013). *Supporting Future Scientists: Predicting Minority Student Participation in the STEM Opportunity Structure in Higher Education* [National Institute of General Medical Sciences, NIH Grant Numbers 1 R01 GMO71968-01 and R01 GMO71968-05 as well as the National Science Foundation, NSF Grant Number 0757076]. The National Association for Research in Science Teaching. http://www.heri.ucla.edu/nih/downloads/NARST2013-Predicting-Minority-Student-Participation-in-STEM.pdf

Filsecker, M. (2011). *The Bachelor Curriculum in the USA Higher Education System: How is it Different?* [Virginia Tech]. https://blogs.lt.vt.edu/greatponderingprocess/files/2012/03/Filsecker-2011_Bachelor-Curriculum-USA-HE.pdf

Flaherty, C. (2021). Defying trends, Ohio State and Syracuse will hire many new tenure-track faculty members. *Inside HigherEd*. https://www.insidehighered.com/news/2021/03/24/defying-trends-ohio-state-and-syracuse-will-hire-many-new-tenure-track-faculty

Ford, K. S., Choi, J., & Baker, D. P. (2020). What if a college major isn't enough? Cognitive skills and the relationship between US college majors and earnings. *Industry and Higher Education*, *34*(6), 410–420. https://doi.org/10.1177/0950422220908187

Fouad, N. A., & Santana, M. C. (2017). SCCT and underrepresented populations in STEM fields: Moving the needle. *Journal of Career Assessment*, *25*(1), 24–39.

Fry, R., & Cilluffo, A. (2019). *A rising share of undergraduates are from poor families, especially at less selective colleges*. Pew Research Center. https://www.pewresearch.org/social-trends/2019/05/22/a-rising-share-of-undergraduates-are-from-poor-families-especially-at-less-selective-colleges/

Ganley, C. M., George, C. E., Cimpian, J. R., & Makowski, M. B. (2018). Gender Equity in College Majors: Looking Beyond the STEM/Non-STEM Dichotomy for Answers Regarding Female Participation. *American Educational Research Journal*, *55*(3), 453–487. https://doi.org/10.3102/0002831217740221

George-Jackson, C. E., Rincon, B., & Martinez, M. G. (2012). Low-Income Engineering Students: Considering Financial Aid and Differential Tuition. *Journal of Student Financial Aid*, *42*(2). https://ir.library.louisville.edu/cgi/viewcontent.cgi?article=1009&context=jsfa

Glennie, E., Mason, M., Dalton, B., & Edmunds, J. (2019). Preparing students for STEM college and careers: The influence of redesigned high schools in North Carolina. *The High School Journal*, *102*(3), 228–257.

Godwin, A. (2016). *The development of a measure of engineering identity*. ASEE Annual Conference & Exposition.

Godwin, A., Potvin, G., & Hazari, Z. (2013). *The development of critical engineering agency, identity, and the impact on engineering career choices*. 1–14.

Gottlieb, J. J. (2018). STEM career aspirations in Black, Hispanic, and White ninth-grade students. *Journal of Research in Science Teaching*, *55*(10), 1365–1392.

Hoff, K., Van Egdom, D., Napolitano, C., Hanna, A., & Rounds, J. (2022). Dream Jobs and Employment Realities: How Adolescents' Career Aspirations Compare to Labor Demands and Automation Risks. *Journal of Career Assessment*, *30*(1), 134–156. https://doi.org/10.1177/10690727211026183

Hughes, B. E. (2018). Coming out in STEM: Factors affecting retention of sexual minority STEM students. *Science Advances*, *4*(3), eaao6373.

Huneeus, F., Miller, C., Neilson, C., & Zimmerman, S. (2021). Firm Sorting, College Major, and the Gender Earnings Gap. *The Working Papers Series of the Central Bank of Chile*, *917*, 43.

Ingels, J. S., Pratt, J. D., Herget, R. D., Burns, J. L., Dever, A. J., Ottem, R., Rogers, E. J., Jin, Y., Leinwand, S., & LoGerfo, L. (2011). *High School Longitudinal Study of 2009 (HSLS:09): Base-Year Data File Documentation* (NCES 2011-328; p. 854). US Department of Education. https://nces.ed.gov/surveys/hsls09/pdf/2011328_1.pdf#page=149&zoom=100,93,96

Ingels, S. J., Herget, D., Pratt, D. J., Dever, J., Copello, E., Leinwand, S., & LoGerfo, L. (2011). *High School Longitudinal Study of 2009 (HSLS:09) Base-Year Field Test Report: Working Paper Series* (p. 884) [Working Paper Series]. US Department of Education, IES-NCES. https://files.eric.ed.gov/fulltext/ED512407.pdf

Ingels, S., Pratt, D., Herget, D., Bryan, M., Fritch, L., Ottem, R., Rogers, J., & Wilson, D. (2015). *High School Longitudinal Study of 2009 (HSLS: 09) 2013 Update and High School Transcript Data File Documentation (NCES 2015-036). US Department of Education. Washington, DC: National Center for Education Statistics.*

Itauma, I. (2019). *Pre-College Factors That Predict Intentions of Minority Females to Enroll in College STEM Programs*. Keiser University.

Jelks, S. M. R., & Crain, A. M. (2020). Sticking with STEM: Understanding STEM Career Persistence among STEM Bachelor's Degree Holders. *The Journal of Higher Education*, *91*(5), 805–831. https://doi.org/10.1080/00221546.2019.1700477

Jewett, E. C. (2019). *AP STEM Course-taking and College STEM Major Selection: An Examination of the Relationship and How It Differs by Gender and Race/Ethnicity* [Seton Hall University]. https://scholarship.shu.edu/cgi/viewcontent.cgi?article=3710&context=dissertations

Jiang, S., Simpkins, S. D., & Eccles, J. S. (2020). Individuals' math and science motivation and their subsequent STEM choices and achievement in high school and college: A longitudinal study of gender and college generation status differences. *Developmental Psychology*, *56*(11), 2137.

Johnson, C. C., Mohr-Schroeder, M. J., Moore, T. J., & English, L. D. (2020a). *Handbook of Research on STEM Education*. Routledge.

Jones, M. C., Srite, M., Chandrasekaran, R., Iyer, L. S., Kayworth, T., & Thatcher, J. B. (2019). Getting Information Systems Programs Classified as STEM: A US-based Perspective from an AIS Task Force Study and Panel Discussion. *Communications of the Association for Information Systems*, *44*(1), 23.

Jones, N., Marks, R., Ramirez, R., & Rios-Vargas, M. (2021). *2020 Census Illuminates Racial and Ethnic Composition of the Country*. US Census Bureau. https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html

Kim, Y. K., & Sax, L. J. (2017). The Impact of College Students' Interactions with Faculty: A Review of General and Conditional Effects. In M. B. Paulsen (Ed.), *Higher Education: Handbook of Theory and Research: Published under the Sponsorship of the Association for Institutional Research (AIR) and the Association for the Study of Higher Education (ASHE)* (pp. 85–139). Springer International Publishing. https://doi.org/10.1007/978-3-319-48983-4_3

Kimball, E. W., Wells, R. S., Ostiguy, B. J., Manly, C. A., & Lauterbach, A. A. (2016). Students with disabilities in higher education: A review of the literature and an agenda for future research. In *Higher education: Handbook of theory and research* (pp. 91–156). Springer.

Kitchen, J. A., Sadler, P., & Sonnert, G. (2018). The Impact of Summer Bridge Programs on College Students' STEM Career Aspirations. *Journal of College Student Development*, *59*(6), 698–715.

Kurban, E. R., & Cabrera, A. F. (2020). Building Readiness and Intention Towards STEM Fields of Study: Using HSLS:09 and SEM to Examine This Complex Process among High School Students. *The Journal of Higher Education*, *91*(4), 620–650. https://doi.org/10.1080/00221546.2019.1681348

Lee, H. R., von Keyserlingk, L., Arum, R., & Eccles, J. S. (2021). Why Do They Enroll in This Course? Undergraduates' Course Choice From a Motivational Perspective. *Frontiers in Education*, *6*, 50. https://doi.org/10.3389/feduc.2021.641254

Lee, J., Weis, L., Liu, K., & Kang, C. (2017). Which type of high school maximizes students' college match? Unequal pathways to postsecondary destinations for students from varying high school settings. *The Journal of Higher Education*, *88*(4), 529–560.

Lehman, K. J., Sax, L. J., & Zimmerman, H. B. (2016). Women planning to major in computer science: Who are they and what makes them unique? *Computer Science Education*, *26*(4), 277–298. https://doi.org/10.1080/08993408.2016.1271536

Lent, R. W., & Brown, D. S. (2019). Social cognitive career theory at 25: Empirical status of the interest, choice, and performance models. *Journal of Vocational Behavior*, *115*, 103316.

Lent, R. W., Brown, S. D., & Hackett, G. Journal of vocational behavior. (1994). *Toward a unifying social cognitive theory of career and academic interest, choice, and performance* (No. 1). *45*(1), 79–122.

Lent, R. W., Miller, M. J., Smith, P. E., Watford, B. A., Hui, K., & Lim, R. H. (2015). Social cognitive model of adjustment to engineering majors: Longitudinal test across gender and race/ethnicity. *Journal of Vocational Behavior*, *86*, 77–85.

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., & Truong, N. N. %J J. of counseling psychology. (2018). *Predictors of science, technology, engineering, and mathematics choice options: A meta-analytic path analysis of the social–cognitive choice model by gender and race/ethnicity*. *65*(1), 17.

Li, S. (2019). *Big Frog in a Small Pond: Undermatching Status, College Major, and Their Influence on Early Career Earnings* [Ph.D., University of California, Los Angeles]. https://www.proquest.com/docview/2388055534/abstract/24B7BBF2A54A4994PQ/1

Lichtenberger, E., & George-Jackson, C. (2012). Predicting High School Students' Interest in Majoring in a STEM Field: Insight into High School Students' Postsecondary Plans. *Journal of Career and Technical Education*, *28*(1). https://doi.org/10.21061/jcte.v28i1.571

Litzler, E., Samuelson, C. C., & Lorah, J. A. (2014). Breaking it Down: Engineering Student STEM Confidence at the Intersection of Race/Ethnicity and Gender. *Research in Higher Education*, *55*(8), 810–832. https://doi.org/10.1007/s11162-014-9333-z

Manly, C. A., Wells, R. S., & Kommers, S. (2018). The influence of STEM definitions for research on women's college attainment. *International Journal of STEM Education*, *5*(1), 45. https://doi.org/10.1186/s40594-018-0144-1

Marade, A. A., & Brinthaupt, T. M. (2018). Good and Bad Reasons For Changing a College Major: A Comparison of Student and Faculty Views. *Education*, *138*(4), 323–337.

Marsh, D. D. (2020b). *Am I a Math or Science Person? How High School Students' Attitudes Towards Mathematics and Science Influence Their Decision to Major in STEM* [University of New Hampshire]. https://www.proquest.com/docview/2423452509?pq-origsite=gscholar&fromopenview=true

Mau, W. J., & Li, J. (2018). Factors influencing STEM career aspirations of underrepresented high school students. *The Career Development Quarterly*, *66*(3), 246–258.

Mau, W.-C. J. (2016). Characteristics of US Students That Pursued a STEM Major and Factors That Predicted Their Persistence in Degree Completion. *Universal Journal of Educational Research*, *4*(6), 1495–1500.

McQuillen, G. (2016). *Conceptually- and Procedurally-Oriented Mathematics Instruction, Achievement, and Equity: Results From the High School Longitudinal Study (HSLS:2009)*. https://dash.harvard.edu/handle/1/33797241

Mehmetoglu, M., & Jakobsen, T. G. (2017). *Applied statistics using Stata: A guide for the social sciences* (1st ed.). Sage.

Mills, D. C. (2019). *Examining the Associations Between Parental Socialization and Science Motivation in Rural African American Adolescents: A Study of Expectancy-Value Theory*.

Minaya, V. (2018). *Revisiting the Gender and Race Disparities in STEM Graduation and Major Choice: The Role of Post-Enrollment Factors*. 50.

Moakler Jr, M. W., & Kim, M. M. (2014). College major choice in STEM: Revisiting confidence and demographic factors. *The Career Development Quarterly*, *62*(2), 128–142.

National Academies of Sciences, Engineering, and Medicine. (2016). *Barriers and Opportunities of 2-year and 4-year STEM degrees* (Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic change to support diverse student pathways). National Academies of Sciences, Engineering, and Medicine. https://www.nap.edu/catalog/21739/barriers-and-opportunities-for-2-year-and-4-year-stem-degrees

National Center for Women & Information Technology. (2021). *Women and Information Technology by the Nnumbers* [Www.ncwit.org]. National Center for Women & Information Technology. https://wpassets.ncwit.org/wp-content/uploads/2021/12/13163828/1209_2021_BTN_FullSize.pdf

National Science Foundation. (2019). *Women, Minorities, and Persons with Disabilities in Science and Engineering*. https://ncses.nsf.gov/pubs/nsf19304/digest/about-this-report

Niu, L. (2017). Family Socioeconomic Status and Choice of Stem Major in College: An Analysis of a National Sample. *College Student Journal*, *51*(2), 298–312.

Nix, S., & Perez-Felkner, L. (2019). Difficulty Orientations, Gender, and Race/Ethnicity: An Intersectional Analysis of Pathways to STEM Degrees. *Social Sciences*, *8*(2), 43. https://doi.org/10.3390/socsci8020043

Olson, S., & Rapporteurs, L. J. B. (2012). *Community Colleges in the Evolving STEM Education Landscape: Summary of a Summit* [Summary of a Summit]. National Research Council and National Academy of Engineering. https://doi.org/10.17226/13399

Orr, M. K., Ohland, M. W., Lord, S. M., & Layton, R. A. (2020). Comparing the Multiple-Institution Database for Investigating Engineering Longitudinal Development with a National Dataset from the United States. *International Journal of Engineering Education*, *36*(4), 1321–1332.

Oseguera, L., & Hwang, J. (2014). *Using Large Data Sets to Study College Education Trajectories*. https://onlinelibrary.wiley.com/doi/full/10.1002/ir.20044

Oseguera, L., Rios, J. D. L., Park, H. J., Aparicio, E. M., & Rao, S. (2020). Understanding Who Stays in a STEM Scholar Program for Underrepresented Students: High-Achieving Scholars and Short-Term Program Retention. *Journal of College Student Retention: Research, Theory & Practice*, 1521025120950693.

Park, J. J., Kim, Y. K., Salazar, C., & Hayes, S. (2020). Student–Faculty Interaction and Discrimination from Faculty in STEM: The Link with Retention. *Research in Higher Education*, *61*(3), 330–356. https://doi.org/10.1007/s11162-019-09564-w

Patnaik, A., Wiswall, M. J., & Zafar, B. (2020). *College Majors* (No. w27645). National Bureau of Economic Research. https://doi.org/10.3386/w27645

Pérez-Sindín, X. S. (2017). Secondary Data: Sources, advantages and disadvantages. In *Secondary Data: Sources, advantages and disadvantages.* (pp. 1577–1579). SAGE Publications, Inc.

Perlow, E. L., Wells, R. S., Hedayet, M., Xia, J., MacLean, H., Ding, E., & McCall, A. (2021). Hidden Inequality: Financial Aid Information Available to College Students with Disabilities Attending Public Four-Year Institutions. *Journal of Student Financial Aid*, *50*(3). https://ir.library.louisville.edu/cgi/viewcontent.cgi?article=1709&context=jsfa

Price, H. E. (2020). The College Preparatory Pipeline: Disparate Stages in Academic Opportunities. *American Educational Research Journal*, 0002831220969138.

Pu, S., Yan, Y., & Zhang, L. (2021). Do Peers Affect Undergraduates' Decisions to Switch Majors? *Educational Researcher*, 0013189X2110235. https://doi.org/10.3102/0013189X211023514

Quadlin, N. (2020). From Major Preferences to Major Choices: Gender and Logics of Major Choice. *Sociology of Education*, *93*(2), 91–109. https://doi.org/10.1177/0038040719887971

Rangel, V. S., Vaval, L., & Bowers, A. (2020). Investigating underrepresented and first-generation college students' science and math motivational beliefs: A nationally representative study using latent profile analysis. *Science Education*, *104*(6), 1041–1070. https://doi.org/10.1002/sce.21593

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

Rincon, B. E., & George-Jackson, C. E. (2016a). Examining Department Climate for Women in Engineering: The Role of STEM Interventions. *Journal of College Student Development*, *57*(6), 742–747. https://doi.org/10.1353/csd.2016.0072

Rincon, B. E., & George-Jackson, C. E. (2016b). STEM intervention programs: Funding practices and challenges. *Studies in Higher Education*, *41*(3), 429–444.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher*, *39*(2), 142–151. https://doi.org/10.3102/0013189X10363170

Schwintz, S. L. (2019). *The Correlation between Financial Support and College Student Attrition: A Quantitative Study*. https://ttu-ir.tdl.org/bitstream/handle/2346/85352/SCHWINTZ-DISSERTATION-2019.pdf?sequence=1

Sheu, H.-B., Lent, R. W., Miller, M. J., Penn, L. T., Cusick, M. E., & Truong, N. N. %J J. of V. B. (2018). *Sources of self-efficacy and outcome expectations in science, technology, engineering, and mathematics domains: A meta-analysis*. *109*, 118–136.

Shewach, O. R., McNeal, K. D., Kuncel, N. R., & Sackett, P. R. (2019). Bunny Hill or Black Diamond: Differences in Advanced Course-Taking in College as a Function of Cognitive Ability and High School GPA. *Educational Measurement: Issues and Practice*, *38*(1), 25–35.

Stange, K. (2015). Differential Pricing in Undergraduate Education: Effects on Degree Production by Field. *Journal of Policy Analysis and Management*, *34*(1), 107–135. https://doi.org/10.1002/pam.21803

Stearns, E., Bottia, M. C., Giersch, J., Mickelson, R. A., Moller, S., Jha, N., & Dancy, M. (2020). Do Relative Advantages in STEM Grades Explain the Gender Gap in Selection of a STEM Major in College? A Multimethod Answer. *American Educational Research Journal*, *57*(1), 218–257. https://doi.org/10.3102/0002831219853533

Stewart, S., Lim, D. H., & Kim, J. (2015). Factors influencing college persistence for first-time students. *Journal of Developmental Education*, 12–20.

Stoddard, C., Urban, C., & Schmeiser, M. D. (2018). College Financing Choices and Academic Performance. *Journal of Consumer Affairs*, *52*(3), 540–561. https://doi.org/10.1111/joca.12175

Thomas, S. L., & Heck, R. H. (2001). Analysis of Large-Scale Secondary Data in Higher Education Research: Potential Perils Associated with Complex Sampling Designs. *Research in Higher Education*, *42*(5), 24.

Thompson, J. A. (2020). *High School Math Course Enrollment among High Math Ability Students: Gender, Race/Ethnicity, Motivation, and Context* [University of Maryland, Baltimore County]. https://www.proquest.com/docview/2408555741?pq-origsite=gscholar&fromopenview=true

Tran, Q. N., Meyerink, M., Aylward, A., & Luo, F. (2021). College Enrollment and STEM Major Choice in a Rural State: A Statewide Examination of Recent High School Cohorts. *Theory & Practice in Rural Education*, *11*(1), Article 1. https://doi.org/10.3776/tpre.2021.v11n1p40-59

Trusty, J., & Niles, S. G. (2003). High-School Math Courses and Completion of the Bachelor's Degree. *Professional School Counseling*, *7*(2), 99–107.

U.S. Bureau of Labor Statistics. (2020). *Employment in STEM occupations 2019 and projected 2029*. https://www.bls.gov/emp/tables/stem-employment.htm

US Census Bureau. (2019). *Quick Facts United States*. https://www.census.gov/quickfacts/fact/table/US/PST045219

U.S. Department of Education, N. C. for E. S. (2016). *Status and Trends in the Education of Racial and Ethnic Groups*. https://nces.ed.gov/programs/raceindicators/indicator_REG.asp

Valliant, R., Li, J., & Liao, D. (2009). *Regression Diagnostics for Survey Data*. 32.

Vu, P., Harshbarger, D., Crow, S., & Henderson, S. (2019). Why STEM? Factors that Influence Gifted Students' Choice of College Majors. *International Journal of Technology in Education and Science*, *3*(2), 63–71.

Wang, M.-T., & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educational Psychology Review*, *29*(1), 119–140.

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.

Wang, X., Lee, Y., & Wickersham, K. (2019). Exploring the Relationship Between Longitudinal Course-Taking Patterns and In-State Transfer Into STEM Fields of Study. *The Journal of Higher Education*, *90*(2), 272–297. https://doi.org/10.1080/00221546.2018.1488210

White House. (2009). *Https://obamawhitehouse.archives.gov/the-press-office/president-obama-launches-educate-innovate-campaign-excellence-science-technology-en#:~:text=President%20Obama%20today%20launched%20the,achievement%20over%20the%20next%20decade.* White House Press Secretary. https://obamawhitehouse.archives.gov/the-press-office/president-obama-launches-educate-innovate-campaign-excellence-science-technology-en#:~:text=President%20Obama%20today%20launched%20the,achievement%20over%20the%20next%20decade.

Wicht, A., Miyamoto, A., & Lechner, C. M. (2021). Are Girls More Ambitious Than Boys? Vocational Interests Partly Explain Gender Differences in Occupational Aspirations. *Journal of Career Development*, 0894845321991665. https://doi.org/10.1177/0894845321991665

Wigfield, A., & Eccles, J. S. (2020). 35 years of research on students' subjective task values and motivation: A look back and a look forward. In *Advances in motivation science* (Vol. 7, pp. 161–198). Elsevier.

Wigfield, A., Muenks, K., & Eccles, J. S. (2021). Achievement Motivation: What We Know and Where We Are Going. *Annual Review of Developmental Psychology*, *3*(1), 87–111. https://doi.org/10.1146/annurev-devpsych-050720-103500

Williams, R. (2021). *Analyzing Complex Survey Data: Some key issues to be aware of*. Https://Www3.Nd.Edu/. https://www3.nd.edu/~rwilliam/xsoc73994/SvyCautionsX.pdf

Wolniak, G. C. (2016). Examining STEM Bachelor's Degree Completion for Students With Differing Propensities at College Entry. *Journal of College Student Retention: Research, Theory & Practice*, *18*(3), 287–309. https://doi.org/10.1177/1521025115622782

Wright, A. L. (2019). *What is in a Major? Habitus, class, gender, and major choice among first-generation college students*. The Ohio State University.

Wu, Q., Dbouk, W., Hasan, I., Kobeissi, N., & Zheng, L. (2021). Does gender affect innovation? Evidence from female chief technology officers. *Research Policy*, *50*(9), 104327. https://doi.org/10.1016/j.respol.2021.104327

Wyner, J., Deane, K. C., Jenkins, D., & Fink, J. (2016). The Transfer Playbook: Essential Practices for Two- and Four-Year Colleges. In *Aspen Institute*. Aspen Institute. https://eric.ed.gov/?id=ED565894

Zafar, B. (2013). College Major Choice and the Gender Gap. *Journal of Human Resources*, *48*(3), 545–595. https://doi.org/10.3368/jhr.48.3.545

Zhang, Y., & Ozuna, T. (2015). Pathways to engineering: The validation experiences of transfer students. *Publication Cover Community College Journal of Research and Practice*, *39*(4), 355–365.

APPENDICES

APPEDIX A. List of variable names

Outcome variable	HSLS:09 Variable Name(s)	Variable	Description
Outcome variable STEM major	X4RFDGMJSTEM	Binary	Reference degree's first major is an NSF STEM field STEM = Science, engineering, and math 1 = yes and 0 = no
Demographic factors			
Gender	X1SEX	Binary	0 = male 1 = female
Race/ethnicity	X1RACE	Categorical	0 = White (*reference group*) 1 = Amer. Indian/Alaska Native and Native Hawaiian/Pacific Islander 2 = Asian 3 = Black/African-American 4 = Hispanic 5 = More than one race
SES	X1SES	Continuous	Composited variable, including parent/guardians' education, occupation, and family income
Math Attainment value	X1MTHID (Student scale: Cronbach's α: 0.84)	Continuous	* You see yourself as a math person (S1MPERSON1) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree * Others see you as a math person (S1MPERSON2) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree
Math Utility value	X1MTHUTI (Student scale: Cronbach's α: 0.78)	Continuous	How much do you agree or disagree with the following statements about the usefulness of

			your [fall 2009 math] course? What students learn in this course... "an advanced math course such as pre-calculus or calculus", "Statistics or Probability", "Algebra II", "Trigonometry", "Analytic Geometry", "Geometry", "Algebra I", "Integrated Math II or above", "Integrated Math I", "Pre-algebra", "Review or Remedial Math" • is useful for everyday life (S1MUSELIFE) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • will be useful for college (S1MUSECLG) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • will be useful for a future career (S1MUSEJOB) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree
Math efficacy	X1MTHEFF (Student scale: Cronbach's α: 0.90)	Continuous	How much do you agree or disagree with the following statements about your [fall 2009 math] course? • You are confident that you can do an excellent job on tests in this course (S1MTESTS) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree

				• You are certain that you can understand the most difficult material presented in the textbook used in this course (S1MTEXTBOOK) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • You are certain that you can master the skills being taught in this course (S1MSKILLS) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • You are confident that you can do an excellent job on assignments in this course (S1MASSEXCL) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree
Science Motivation and Confidence				
Science value	Attainment	X1SCIID (Student scale: Cronbach's α: 0.83)	Continuous	How much do you agree or disagree with the following statements? You see yourself as a science person (S1SPERSON1) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree Others see you as a science person (S1SPERSON2) 1=Strongly agree 2=Agree 3=Disagree

			4=Strongly disagree
Science Utility value	X1SCIUTI (Student scale: Cronbach's α: 0.75)	Continuous	How much do you agree or disagree with the following statements about the usefulness of your [fall 2009 science] course? What students learn in this course... "Advanced Physics", "Advanced Chemistry", "Advanced Biology", "Anatomy or Physiology", "Environmental Science", "Integrated Science II or above", "Integrated Science I", "Principles of Technology", "Physics I", "Chemistry I", "Biology I", "a biological sciences course", "Earth Science", "an earth or environmental science course", "Life Science", "Physical Science", "a physical science course", "General Science" • is useful for everyday life (S1SUSELIFE) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • will be useful for college (S1SUSECLG) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • will be useful for a future career (S1SUSEJOB) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree
Science self-efficacy	X1SCIEFF (Student scale: Cronbach's α: 0.88)	Continuous	How much do you agree or disagree with the following statements about your [fall 2009 science] course?

			• You are confident that you can do an excellent job on tests in this course (S1STESTS) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • You are certain you can understand the most difficult material presented in the textbook used in this course (S1STEXTBOOK) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • You are certain you can master the skills being taught in this course (S1SSKILLS) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree • You are confident that you can do an excellent job on assignments in this course (S1SASSEXCL) 1=Strongly agree 2=Agree 3=Disagree 4=Strongly disagree
High school GPA	X3TGPATOT	Continuous	High school GPA
Enrolled Degree’s program level	X3PROGLEVEL	Categorical	0 University transfer Associate's degree program 1 Bachelor's degree program

College STEM credits earned in college	X5STEM3ERN	Continuous	Postsecondary Transcript: STEM (using NCES grant definition of STEM): known credits earned
College STEM GPA	X5STEM3GPA	Continuous	Postsecondary Transcript: STEM (using NCES grant definition of STEM): known GPA
Collegiate Supports			
Merits-aid	X5PFYMERITAID	Continuous	Total merit-only grants at primary first year institution in USD. the sum of state merit-only grants and scholarships (X5PFYSTMERIT) and institutional merit-only grants and scholarships, including athletic scholarships (X5PFYINSMERIT)
Needs-based aid	X5PFYNEEDAID	Continuous	Total need-based grants at primary first year institution in USD. The sum of federal Pell Grants (X5PFYPELLAMT, from NSLDS:17), Federal Supplemental Educational Opportunity Grant (SEOG) awards (X5PFYSEOGAMT, from the 2017 Student Records Collection), state need-based grants (X5PFYSTATNEED, from the 2017 Student Records Collection), and institutional need-based grants (X5PFYINSTNEED, from the 2017 Student Records collection) received while enrolled at the primary institution during the first academic year attended postsecondary education after high school
Research participation with faculty	S4RESEARCH	Binary	Did you participate in any of the following as a part of your college or trade school education? 0 = No , 1= Yes
Academic support services used	S4SRVACAD	Binary	Service used: academic support services (for example, tutoring or writing centers) (Visiting, emailing, or in any way communicating with and

			receiving information or help from a school office or department that offers a particular service counts as use of that service.) 0 = No , 1 = Yes
Career support services used	S4SRVCAREER	Binary	Service used: career planning or job placement services 0 = No , 1 = Yes
Personal Circumstances			
Work vs. Academic performance	S4WRKINTERFERE		Work schedule interfered with academic performance in college 0 = No , 1 = Yes
Money concerns in 2015	S4EVERWRYMNY	Binary	In calendar year 2015, did you ever...worry about having enough money for regular expenses? 0 = No, 1 = Yes
Disability/Special need	X4DISABLED	Binary	This variable indicates if the respondent ever had disability or special need 0 = No, 1 = Yes This variable indicates if any of the following had ever been true for the respondent; he/she 1) had a serious difficulty concentrating, remembering, or making decisions (see also S4DIFCONC), 2) had been told by a health or education professional that he/she had ADHD or ADD (Attention Deficit Hyperactivity Disorder or Attention Deficit Disorder) (see also S4ADHD), 3) had a learning disability (see also S4LEARNDISBL), 4) was deaf or had a serious difficulty hearing (see also S4DEAF), 5) was blind or had a serious difficulty seeing (see also S4BLIND), or 6) had any other disability or special need (see also S4OTHDISBL).

APPENDIX B. STEM Policy in the United States (Johnson et al., 2020)

Year/Org.	Abbreviations	Description
1965	ESEA	Elementary and Secondary Education Act
Early 1990s	STEM education	Initially SMET, changed to STEM in 2001 by NSF
2002	NCLB Act	No Child Left Behind Act (K-12 education)
2005		*Rising Above the Gathering Storm* report (National Academies)
2007	America COMPETES Act	America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (K-12)
2009 (after the Great Recession)	ARRA	American Recovery and Reinvestment Act, including the Race to the Top (RTTT) competitive grant program ($4.35bn) impacts of RTTT in selected states increased college readiness outcomes for STEM schools (p.402)
2010	America COMPETES Act	More focus on postsecondary education
2010		President’s Council of Advisors in Science and Technology (PCAST)
2010, NSF	STEM definition	Science, technology, engineering, math, psychology, social sciences, physical and life sciences
Dept. of Homeland Security (DHS), US Immigration and Customs Enforcement (ICE)	STEM definition	STEM exclude most social sciences
Standard Occupational Classification (SOC) Policy Committee	STEM definition	STEM occupations into two broad domains: - Science, Engineering, Mathematics, and IT - Science and Engineering – related occupations
2009, 2013		White House’s Educate to Innovate

2015, 2017	America COMPETES Act	The act was never fully funded and most of the new STEM education programs associated with the act were either “partially implemented or not implemented at all” (Congressional Research Service)
2015	ESSA	ESEA was reauthorized under the auspices of ESSA. ESSA shifts responsibility for educational accountability to the states. Impact of ESSA legislation, especially for STEM, was limited.
2015	STEM Education Act	No authorized funding
2013, 2018		5-year strategic plan to support STEM education issued by the National Science and Technology Council (NSTC) NSTC’s subcommittee CoSTEM (Committee on STEM Education)
2018		Indicators for Monitoring Undergraduate STEM Education (National Academy of Sciences)
2018		STEM initiatives profile published by the Education Commission of the States (ECS) • 38 states’ policies provide financial incentives in STEM teach recruitment and 23 provided STEM professional development for high-school teachers Student-level policies: • 8 states’ policies support student STEM mentoring and internships • 12 support afterschool STEM programs • 12 support programs targeting underrepresented groups in STEM
2018	Perkins Act	Passed in 1984 for Vocational and Technical Education

	Reauthorized in 1998, 2006. Reauthorized in 2018 as "Strengthening CTE for the 21st Century Act".

APPENDIX C. Comparisons Between Excluded Sample and Analytic Sample (Transcripts data)

Variable	Entry-college stage Excluded sample (n=10,619)				Third-year college stage Analytic sample (n=7,458)			
	n	Mean/%	SE	Min-Max	n	Mean /%	SE	Min-Max
Dependent variable	10,619			0-1	7,458			0-1
STEM		20%	.02			22%	0.10	
Non-STEM		68%	.09			78%	0.10	
Missing data		12%		(-9/-1)				
Independent variables								
Enrolled degree program level								
Bachelor's						60%	.01	
Associate's						40%	.01	
Gender	10,619			0-1	7,458			0-1
Male (0)		48%	.01			47%	.01	
Female (1)		52%	.01			53%	.01	
Race				1-8				0-4
Asian		5%	.01			4%	.01	
Black		11%	.01			10%	.01	
Hispanic		17%	.01			16%	.02	
More than one race		8%	.01			8%	.01	
White (reference)		59%	.04			61%	.02	
AI – HI/PI		.04%	.00					
SES	10,619	.23	.03	(-1.82) – 2.57	7,458	.27	.04	(-1.75) -2.57
Math attainment	10,554	.19	.02	(-1.73) – 1.76	7,401	.22	.02	(-1.73) -1.76
Math utility	10,048	.03	.02	(-3.51) – 1.31	6,962	.01	.02	(-3.51) -1.31
Math self-efficacy	10,049	.18	.02	(-2.92) – 1.62	6,965	.20	.03	(-2.92) - 1.62
Science attainment	10,562	.20	.02	(-1.57) – 2.15	7,407	.23	.03	(-1.57) - 2.15
Science utility	9,676	.05	.03	(-3.1) – 1.69	6,633	.04	.03	(-3.1) – 1.69
Science self-efficacy	9,670	.14	.02	(-2.91) – 1.83	6,634	.16	.02	(-2.91) - 1.83
HS_GPA	10.366	3.07	.02	.25 – 4.0	7,233	3.14	.02	.25 – 4.0
College-STEM credits earned	10,559	20.04	.56	0- 115	7,413	21.31	.68	0- 115

Table 2 Continued

College STEM GPA	10,042	2.50	.02	0 - 4	7,010	2.57	.04	0 - 4
Merit aid [$0 - 38,607]	10,619			0-1	7,458			0-1
No (0)		76%	.01			61%	.02	
Yes (1)		24%	.01			39%	.02	
Need aid [$0-50,340]	10,619			0-1	7,458			0-1
No (0)		64%	.02			47%	.02	
Yes (1)		36%	.02			53%	.02	
Research with faculty	9,965			0-1	7,056			0-1
No (0)		83%	.01			82%	.01	
Yes (1)		17%	.01			18%	.01	
Academic services used	10,006			0-1	7,098			0-1
No (0)		50%	.02			48%	.01	
Yes (1)		50%	.02			52%	.01	
Career service used	10,006			0-1	7,098			0-1
No (0)		69%	.01			68%	.01	
Yes (1)		31%	.01			32%	.01	
Work schedule & academic performance interference	8,609			0-1	5,819			0-1
No (0)		54%	.02			55%	.01	
Yes (1)		46%	.02			45%	.01	
Worry about money expenses in 2015	10,093			0-1	7,024			0-1
No (0)		44%	.01			44%	.02	
Yes (1)		56%	.01			56%	.02	
Disability	10.077			0-1	7,010			0-1
No (0)		70%	.01			71%	.01	
Yes (1)		30%	.01			29%	.01	

APPENDIX D. Descriptive Statistics of Non-missing data and Full-analytic sample

	Non-missing data (n=3,071)				Third-year college stage Analytic sample (n=7,458)			
Variable	n	Mean/%	SE	Min-Max	n	Mean/%	SE	Min-Max
Dependent variable	3,071			0-1	7,458			0-1
STEM		23%	.01			22%	0.10	
Non-STEM		77%	.01			78%	0.10	
Independent variables								
Enrolled degree's program level	3,071			0-1	7,458			0-1
Bachelor's		60%	.02			60%	.01	
Associate's (transfer)		40%	.02			40%	.01	
Gender	3,071			0-1	7,458			0-1
Male (0)		43%	.02			47%	.01	
Female (1)		57%	.02			53%	.01	
Race				0-4				0-4
Asian		5%	.01			4%	.01	
Black		9%	.01			10%	.01	
Hispanic		16%	.02			16%	.02	
More than one race		8%	.01			8%	.01	
White (reference)		62%	.02			61%	.02	
SES	3,071	.28	.04	(-1.75) – 2.57	7,458	.27	.04	(-1.75) -2.57
Math attainment	3,071	.19	.04	(-1.73) – 1.76	7,401	.22	.02	(-1.73) -1.76
Math utility	3,071	.02	.04	(-3.51) – 1.31	6,962	.01	.02	(-3.51) -1.31
Math self-efficacy	3,071	.20	.03	(-2.92) – 1.62	6,965	.20	.03	(-2.92) - 1.62
Science attainment	3,071	.24	.03	(-1.57) – 2.15	7,407	.23	.03	(-1.57) - 2.15
Science utility	3,071	.06	.04	(-3.1) – 1.69	6,633	.04	.03	(-3.1) – 1.69
Science self-efficacy	3,071	.14	.03	(-2.91) – 1.83	6,634	.16	.02	(-2.91) - 1.83
HS_GPA	3,071	3.18	.03	.25 – 4.0	7,233	3.14	.02	.25 – 4.0
College-STEM credits earned	3,071	23.51	.86	0- 115	7,413	21.31	.68	0- 115

Table 3 Continued

College_STEM GPA	3,071	2.60	.05	0 - 4	7,010	2.57	.04	0 - 4
Research with faculty	3,071			0-1	7,056			0-1
No (0)		81%	.02			82%	.01	
Yes (1)		19%	.02			18%	.01	
Academic services used	3,071			0-1	7,098			0-1
No (0)		46%	.02			48%	.01	
Yes (1)		54%	.02			52%	.01	
Career service used	3,071			0-1	7,098			0-1
No (0)		66%	.02			68%	.01	
Yes (1)		34%	.02			32%	.01	
Work schedule & academic performance interference	3,071			0-1	5,819			0-1
No (0)		56%	.01			55%	.01	
Yes (1)		44%	.01			45%	.01	
Money worry about regular expenses in 2015	3,071			0-1	7,024			0-1
No (0)		42%	.02			44%	.02	
Yes (1)		58%	.02			56%	.02	
Disability	3,071			0-1	7,010			0-1
No (0)		72%	.02			71%	.01	
Yes (1)		28%	.02			29%	.01	