ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 19-42

Explaining the Gender Gap in STEM Attainment: Factors from Primary School to STEM Degree Completion

Katherine A. Key
Georgia State University

Tim R. Sass
Georgia State University

We investigate the determinants of high school completion and college attendance, the likelihood of taking science, technology, engineering or math (STEM) courses in the first year of college and the probability of earning a degree in a STEM field. The focus is on women, who tend to be under-represented in STEM fields. Tracking four cohorts of students throughout Florida, women perform nearly as well as men on math achievement tests through high school and are more likely to finish high school and attend college than males. Among college students, however, women are less likely than are men to take courses in the physical sciences in their first year and are less likely to earn a degree in physics or engineering, even after adjusting for pre-college test scores. Gender matching of students and math/science teachers in middle and high school tends to increase the likelihood that female college freshman will take at least one STEM course. However, conditional on first-year coursework, neither gender matching at the secondary or college levels appears to have any effect on the likelihood of completing a major in a STEM field. For all students, having high school math and physics teachers with a degree in math or physics, respectively, (as opposed to education) is associated with a higher likelihood of taking STEM courses as college freshmen.

VERSION: May 2019

Explaining the Gender Gap in STEM Attainment: Factors from Primary School to STEM Degree Completion*

Katherine A. Key
Department of Economics
Georgia State University

Tim R. Sass
Department of Economics
Georgia State University

April 19, 2019

Abstract

We investigate the determinants of high school completion and college attendance, the likelihood of taking science, technology, engineering or math (STEM) courses in the first year of college and the probability of earning a degree in a STEM field. The focus is on women, who tend to be under-represented in STEM fields. Tracking four cohorts of students throughout Florida, women perform nearly as well as men on math achievement tests through high school and are more likely to finish high school and attend college than males. Among college students, however, women are less likely than are men to take courses in the physical sciences in their first year and are less likely to earn a degree in physics or engineering, even after adjusting for pre-college test scores. Gender matching of students and math/science teachers in middle and high school tends to increase the likelihood that female college freshman will take at least one STEM course. However, conditional on first-year coursework, neither gender matching at the secondary or college levels appears to have any effect on the likelihood of completing a major in a STEM field. For all students, having high school math and physics teachers with a degree in math or physics, respectively, (as opposed to education) is associated with a higher likelihood of taking STEM courses as college freshmen.

*We wish to thank the staff of the Florida Department of Education's Education Data Warehouse for their help in obtaining and interpreting the data used in this study. Funding for this research was provided by the Gates and Smith-Richardson Foundations as part of a grant to the Center for the Analysis of Longitudinal Data in Education Research (CALDER). Any opinions or errors are solely attributable to the authors, however.

I. Introduction

There is growing concern that the United States does not produce a sufficient number of students majoring in science, technology, engineering and math (STEM) fields to remain globally competitive. Of particular concern is the underrepresentation of women in STEM fields. Various hypotheses have been put forth to explain why there are relatively few women in STEM areas, including negative peer effects of male students in math and science courses and lack of female instructors as role models.

Most of the existing evidence on enrollment and persistence in STEM majors is based on experiences while in college. However, this may obscure important influences of student preparation and experiences in high school or earlier. The focus on instructors and peers in college is primarily a matter of data availability. Most extant studies have relied on college transcript data, which provide little information about student experiences prior to attending college.

The present study of STEM degree attainment is the first to track individual students from elementary school through the end of college and the first to link individual students to their STEM instructors in middle school, high school and college. It is also the first study to simultaneously consider the impacts of high school instructor demographics and the training of high school teachers on their students' choices of coursework and major in college. The availability of linked K-12 and post-secondary transcript data allow us to determine where gaps emerge between males and females and the relative importance of pre-college and within-college experiences in determining STEM educational attainment. Such information is potentially important in determining where to target interventions designed to promote female participation in STEM fields.

The evidence reveals that women perform nearly as well as men on math achievement tests through high school and are more likely to finish high school and attend college than males. Among college students, however, women are much less likely than men to earn a degree in a STEM field, even after adjusting for pre-college test scores. There is some evidence that gender matching of students and secondary teachers increases the likelihood that women take STEM classes as college freshmen. Likewise, taking high school courses from teachers who possess a baccalaureate degree in math or science (rather than education) is associated with a higher probability of taking STEM courses as a college freshman. However, conditional on first-year college coursework, gender matching of students and college instructors is not associated with completing a degree in a STEM field.

II. Prior Evidence

Spurred by the intense policy interest in the underrepresentation of women in STEM fields, there has been a surge in research on the determinants of entry and persistence in STEM majors. Most of this recent work focuses on major choice conditional upon enrolling in college, with a smaller literature on pre-college determinants of major choice and completion.

A. STEM Persistence Among College Students

Non-experimental studies of instructor-student gender matching in college have produced mixed results. Using data from public colleges and universities in Ohio, Bettinger and Long (2005) find that female students who initially had a female instructor for geology, math or statistics were more likely to take additional courses in the subject than women whose first instructor was male. However, for biology and physics, just the opposite was true; women whose first instructor was female were less likely to take additional courses in the subject than similar students whose first professor was male. Further, they found no effects of instructor gender on the major choices of female college students in STEM fields. Price (2010) analyzes the same Ohio data, but for a longer time period, and finds that female students are less likely to persist as the proportion of their STEM courses taught by female instructors rises. Hoffmann and Oreopoulos (2009) find that gender matching of students and instructors in large first-year undergraduate courses at the University of Toronto has no effect on the likelihood of dropping the course, but does increase the number of additional courses taken in the subject. This later effect does not hold for math and science courses, however, where gender matching actually reduces the number of additional courses taken in the subject. Griffith (2010) finds a zero or negative correlation between percent of female faculty in STEM and female student persistence in a STEM major.

Evidence of positive gender-match effects is much stronger in studies with random assignment of students to professors. Carrell, Page and West (2010) analyze the effects of instructor gender on student achievement, course taking and STEM major choice at the U.S. Air Force Academy, where all cadets take a set of standardized introductory courses and classroom assignments are random. They find substantial positive effects of female instructors on female student's performance in math and science courses at the Academy. Further, high-performing young women, as measured by SAT math scores, are more likely to take additional STEM courses and are more likely to graduate with a degree in a STEM major if they take introductory courses taught by a female instructor. In later work, Mansour, Rees, Rintala and Wozny (2018) find that high-ability female cadets also have higher probabilities of receiving a master's degree in a STEM field and of eventually working in a STEM occupation if they are randomly assigned to a female professor. It is unclear, however, if these strong gender-matching findings are due to the studies' rigorous experimental designs or are a result of unique aspects of the female students attending an elite military academy.

Two studies have investigated the effects of peer composition on female STEM outcomes in college.[1] Kokkelenberg and Sinha (2010) find peer influences at SUNY-Binghampton had mixed effects; increases in the proportion of female students increased the likelihood a woman would earn an "A" in sophomore math courses, but decreased the probability of obtaining an "A" in junior-level math courses and had no significant effect on grades earned in biology courses. Fischer (2017) analyzes gender differences in the effect of peer ability in an introductory general chemistry course (a prerequisite for most STEM majors) on STEM major completion at the University of California - Santa Barbara. Women enrolled in sections with a higher proportion of "on track" students were less likely to graduate with a STEM degree, while men's STEM persistence is not affected by classroom peer composition. The effect is largest for women in the bottom third of the SAT-score distribution.

B. Pre-College Influences on Educational Attainment in STEM Fields

The literature relating pre-college experiences to college outcomes is relatively thin.[2] Two studies analyze the long-run impacts of peer gender composition. Park, Behrman and Choi (2018) exploit the random assignment of students to high schools in Korea to estimate the causal effects

[1] A recent study considers peer influences outside the classroom on STEM attainment. Dennehy and Dasgupta (2017) finds that women who were randomly assigned to female peer mentors had higher rates of retention in engineering majors than those who were not mentored, but there were no differences for female engineering students who were assigned a male mentor.

[2] A number of studies measure the effect of teacher-student gender matching on test scores and beliefs about STEM ability in K-12, but do not gauge impacts on educational outcomes past high school. See Dee (2007); Winters, Haight, Swaim and Pickering (2013); Antecol, Eren and Ozbeklik (2015); Sansone (2017); and Sansone (2019). Another strand of literature investigates the relationship between pre-college test scores and college major selection, but does not consider specific mechanisms like peer effects, course selection or student-teacher matching. See Turner and Bowen (1999), Dickson (2010) and Speer (2017).

of attending same-sex schools. They find that attending an all-girls school has no statistically significant effect on general math test scores, math-science test scores, Korean test scores or English test scores. Further, attending an all-girls high school has no significant effect on interest in science, student expectations or actual choice of a STEM major. Anelli and Peri (2019) investigate the effects of high school gender composition on post-secondary outcomes in Milan. In the Italian system, high school students are randomly grouped into "classes" during their freshman year and maintain these groupings throughout high school. They find no evidence that the share of own-gender peers in high school impacts the probability of choosing a "prevalently male" or "prevalently female" college major for either young men or women.

A pair of studies consider the impact of high school coursework on post-secondary STEM outcomes. Card and Payne (2017) analyze data from a cohort of high school students in the province of Ontario. In Canada, students apply to specific programs at universities within a province and admissions are determined solely by grades in standardized courses taken in the fourth year of high school. Card and Payne find that differences in high school course taking explain only a small proportion of the gender gap in STEM program entry. Rather, the gap is mainly due to higher rates of college going by women, which means that a smaller proportion of female college entrants possess the pre-requisites for STEM programs in college. Bottia, et al. (2015) studies a single cohort of students who graduated from North Carolina public schools in 2004 and enrolled in a public 4-year institution the same year. Their data do not link students and teachers to individual classrooms at either the high school or post-secondary level, so it is not possible to determine the gender of specific instructors. They find that higher proportions of female STEM faculty at the high school level are associated with greater likelihoods that women declare a STEM major and graduate with a STEM degree. These effects are greatest for female students with the strongest math skills.

III. Data

The data for this study come from a variety of sources. The primary source for student-level information is the Florida Department of Education's K-20 Education Data Warehouse (K-20 EDW), an integrated longitudinal database covering all public school students and teachers in the state of Florida. For K-12 students, the K-20 EDW provides demographic information, enrollment and attendance, program participation, disciplinary actions and achievement test scores, beginning in 1995. Florida began testing students statewide in 1997/98, with the introduction of the "Sunshine State Standards" Florida Comprehensive Achievement Test (FCAT-SSS). The FCAT-SSS is a criterion-based exam designed to test for the skills that students are expected to master at each grade level. It is a "high-stakes" test used to determine school grades, student retention in some grades and passage of the 10^{th} grade exam was a requirement for graduation from high school for many years. The FCAT-SSS exam was initially administered to students in selected grades but was later expanded to grades 3-10 in 2000/01. Beginning in 1999/2000, a second test, the FCAT Norm-Referenced Test (FCAT-NRT), was added in each of grades 3-10. The FCAT-NRT was a custom form of the Stanford Achievement Test used throughout the country. No accountability measures were tied to student performance on the FCAT-NRT. Florida stopped administering the FCAT-NRT after 2007/08. The FCAT-SSS exam was replaced with the FCAT 2.0 beginning in 2010/11.

As the name implies, the K-20 EDW also includes records for students enrolled in community colleges or four-year public universities in Florida. The K-20 EDW also contains information on the Florida Resident Assistance Grant (FRAG), a grant available to Florida residents who attend private colleges and universities in the state. Data from the National Student Clearinghouse (NSC), a national database that includes enrollment data from 3,300 colleges throughout the United States, is used to track college attendance outside the state of Florida, as well as any private college enrollment in Florida that the FRAG data do not pick up. Unfortunately, the Florida Department of Education's data-sharing agreement with the NSC expired in the latter part of the 2000s, so we can only reliably track students who attended private colleges and universities within Florida or any postsecondary institution outside of Florida through school year 2006–2007.[3] Enrollment, coursework and degree attainment information are available for all post-secondary students at public institutions in Florida. In addition, demographic information on post-secondary instructors is available as well.

High school graduation status is identified based upon withdrawal information and student award data from the K-20 EDW. While various diploma options exist, including a GED and a special-education diploma, we focus on receipt of a regular high school diploma. Students who withdrew with no intention of returning or exited for other reasons such as non-attendance, court action, joining the military, marriage, pregnancy, and medical problems, but did not later graduate, are counted as dropouts. It is not possible to directly determine the graduation status of students who leave the Florida public school system to attend a home-schooling program, to enroll in a private school or who move out of state.

The analysis sample covers four cohorts of 5th-grade students. Statewide achievement testing for 5th-grade students began in the 1997/98 school year, so the first cohort in the sample are students who attended 5th grade for the first time in 1997/98 and took the FCAT-SSS math exam.

[3] Information on the NSC is available at www.studentclearinghouse.org.

The final cohort is composed of students who were enrolled in 5^{th} grade for the first time in 2000/01. Descriptive statistics for these four cohorts of students are provided in Table 1.

The last year for which we were able to obtain student data is 2012/13. Given that it takes at least three years to progress through middle school and high school completion typically takes four years, this means that each of the four cohorts can be tracked through high school and into the beginning of college. If we allow five years for college completion, then all four cohorts can be tracked through the end of college so long as we restrict the sample to students who do not repeat any grades in middle school. Descriptive statistics for this restricted sample are provided in Table 2. As one would expect, students who follow the normal progression through middle school have higher test scores than those who repeat a grade or drop out before completing high school and thus the test score means in Table 2 are greater than those in Table 1 for the majority of exams for each cohort.

IV. Analysis and Results

A. Descriptive Analysis of Pre-College Outcomes and College Attendance

In order to understand where the "leakages" in the STEM pipeline occur, we begin with a descriptive analysis of achievement differences in math prior to high school entry. Figure 1 illustrates mean test scores in elementary and middle school by gender. Consistent with earlier work by Fryer and Levitt (2010), we find that girls tend to score below boys on math exams in elementary school. However, in contrast to Fryer and Levitt, the differences we find are relatively modest, ranging from less than 0.01 to 0.04 standard deviations. In 8^{th} grade, girls score 0.04

standard deviations above boys, on average, on the FCAT-SSS exam and only about 0.01 less than boys on the FCAT-NRT exam.[4]

The high school experiences of students exhibit patterns to those of the achievement measures in elementary and middle school. Table 3 provides information on exit propensities by gender. As with test scores prior to high school, the differences in drop-out rates between boys and girls are relatively small, with females about one percentage point less likely to drop out of high school.

Even if students complete five years of high school, they may not earn a regular high school diploma. They could receive a GED or (if they are a special education student) a special diploma or certificate of completion. Alternatively, they could remain enrolled, but still not have obtained a diploma within five years. As demonstrated in Table 4, women are about six percentage points more likely to earn a standard high school diploma than are men (about a 10-percent differential).

A final measure of high school performance is achievement test scores. Figure 2 depicts average math test scores by grade from grade 5 through grade 10 for a fixed group of students. Results vary somewhat across exams, but boys maintain about a 0.04 standard deviation in performance across both math exams in 10th grade.

Conditional on earning a regular high school diploma, there are substantial differences in post-secondary educational choices between men and women. As reported in Table 5, women are more likely then men to attend college. The female advantage holds across all types of post-

[4] Interestingly, gender differences for blacks follow a somewhat different pattern than for other racial/ethnic groups, with black girls consistently outperforming black boys in math throughout grades 3-8. While analysis of racial differences in STEM outcomes is beyond the scope of the present paper, a companion paper, Sass (2017) explores factors affecting racial gaps in STEM attainment. Other recent work on racial disparities in STEM and the impacts of same-race teachers includes Gershenson, Holt and Papageorge (2016) Gershenson, Hart Lindsay and Papageorge (2018).

secondary institutions, including both two-year and four-year public in-state institutions, as well in-state private colleges and universities and out-of-state institutions of higher education.

Taken together, the descriptive evidence on test scores and educational attainment prior to college suggest that women should be as likely as men to complete a STEM major while in college. We turn now to analyses of course selection, persistence and major choice among those students who make it to college.

B. Determinants of Initial College Coursework

Table 6 presents probit estimates of the probability of taking one or more STEM courses in the first year of college, conditional upon having earned a regular high school diploma within five years of starting grade 9 and enrolling in a four-year public university within a year of receiving their high school diploma. The first column reports estimates with only student gender and race/ethnicity in the model. Without any other controls, women are about 4 percentage points less likely than men to take at least one STEM course in their freshman year.[5] Estimates reported in the second column indicate that controlling for family income (5^{th}-grade lunch status) and math test scores in grades 5-10 does not substantially alter this finding.[6]

As shown in Table 7, the relationships between STEM course taking and gender vary across science disciplines. Women are less likely than men to take any science course and less likely to take computer science, engineering, math or physics courses during their freshman year. They are more likely than men, however, to take at least one biology course or at least one statistics course during their first year of college.

---

[5] Remedial math courses are excluded.

[6] Much of the influences of income and prior achievement are likely implicitly controlled for by restricting the sample to high school completers who immediately enter a four-year public university in Florida.

Table 8 presents estimates of the determinants of first-year STEM course taking, controlling for the characteristics of middle and high school teachers and peers.[7] The first two specifications exclude high school fixed effects whereas the third and fourth specifications include them and thereby reflect within-high-school comparisons. The fifth specification includes instrumental variables in a probit model as an alternative strategy for addressing potential endogeneity. Due to the incidental parameters problem, probit models with fixed effects may be biased (Greene, 2004). Therefore, we only estimate linear probability models for the fixed-effects specifications. We present estimates from both probit and linear probability models without fixed effects for comparison purposes.

Including high school fixed effects eliminates bias caused by unmeasured time-invariant school characteristics that are correlated with both high school teacher characteristics and with subsequent college course taking decisions. For example, suppose that schools in relatively affluent neighborhoods tend to have teachers with degrees in their subject area and parents in these neighborhoods tend to push their children toward STEM majors in college. What would appear to be an effect of teacher credentials could in fact be caused by unmeasured neighborhood characteristics. Including high school fixed effects would eliminate such potential biases.

The disadvantage of using high school fixed effects is that they may soak up important cross-school variation that would otherwise be captured by the variables of interest. For example, if gender matching of students and teachers in high school promotes later college STEM course taking, we would expect that female students from a school with a predominately female math/science faculty would fare better than would female students attending a school where the

[7] Note that we control for pre-high school math test scores, but not scores in grades 9 and 10, so the high school teacher effects could be working through effects on high school math achievement.

math and science faculty are mostly male. With high school fixed effects, these cross-school differences would be absorbed into the fixed effects, however. Put differently, identification of the effects of faculty identity come solely from within-school changes over time in faculty composition. Given a student typically attends high school for four years and we have only four cohorts of students, within-school variability may be somewhat limited.

For both models without fixed effects, we find very similar positive student-teacher match effects for women. Having half of middle/high school math and science course taught by a female instructor (rather than none) is associated with around a 0.03 increase in the likelihood of taking one or more STEM courses as a college freshman, on par with the estimated gender gap in first-year STEM course taking of -0.0325 (in Table 7). The preparation of math and science teachers matters as well. Taking at least one high school biology, chemistry or math course taught by a teacher with a bachelor's degree in the relevant subject (rather than a math or science education degree or a degree in another science field) is associated with a small but statistically significant increase in the likelihood a student later takes at least one STEM course as a freshman in college. Likely due to the small number of high school instructors with degrees in physics, the estimated impact of having a high school physics teacher who majored in physics is not statistically significant at conventional confidence levels. In contrast to the apparent influences of teachers, having a greater proportion of female students in math and science courses does not appear to be correlated with the likelihood that a young woman will take STEM courses in their first year in college. In fact, the correlation between the fraction of female students in middle and high school math and science courses and STEM course taking during the first year in college is negative.

Many of the significant correlations between the characteristics of high school teachers and college course taking decisions disappear when high school fixed effects are included in the model.

As shown in the third column of Table 8, the positive match effects for women are eliminated when high school fixed effects are employed. Likewise, all of the positive teacher training correlations go away as well. The only statistically significant association that remains is a small negative correlation between having a high school biology course taught by a teacher who majored in biology and later STEM course taking. In an attempt to enhance inter-temporal variability (which is required for identification in the fixed effects model), we limit the sample to large high schools that employed 100 or more unique teachers. Results are presented in the fourth column of Table 8. The female indicator is now negative and statistically significant. In addition, the partial correlation between having a high school chemistry course taught by a teacher with a degree in chemistry and first-year college STEM enrollment is now positive and statistically significant.

As an alternative to the fixed effects models, we also estimate a probit model that deals with potential endogeneity by employing instrumental variables. The advantage of this approach is it allows cross-school variation to be considered while still addressing potential bias from self-selection of high school instructors. Following the strategy employed by Bettinger and Long (2005) and Price (2010), we use average faculty composition variables as instruments for the characteristics of teachers a student has in high school.[8] Thus, for example, the school-wide proportion of biology teachers who possess a degree in biology during a student's high school career serves as an instrument for the indicator that a student was taught biology by a teacher who majored in biology in college. Results from the probit IV estimation are presented in the last column of Table 8. The IV results are qualitatively similar to those from the probit model without high school fixed effects. Gender matching of students and teachers in middle and high school

[8] We calculate school average faculty composition for each student by taking the simple average of school/year/grade faculty composition over each of the school/grade/year combinations for the years each student was enrolled in grades 9-12.

math and science courses is associated with an increase in the likelihood of women taking at least one STEM course as a college freshman. The likelihood of a female student taking one or more STEM classes in their first year in college is positively correlated with being taught high school math or physics by someone who majored in the relevant subject in college. In contrast, it appears that having more female students in middle and high school math and science courses may actually diminish the likelihood of taking a STEM course in the first year of college; a 20 percent increase in the proportion of classroom peers who are women is associated with a 10 percentage point reduction in the probability a female student will take one or more STEM courses during their first year in college.

C. Determinants of Completing a College Degree in a STEM Major

Table 9 presents results of estimating probit equations which predict completion of college majors (conditional on attending a Florida public university immediately after earning a high school degree), with and without controls for pre-college family income and student achievement test scores. The estimates from the model without pre-college controls reveal the expected pattern; women are less likely than men to successfully complete a major in a STEM field. This pattern generally holds across specific STEM majors. The one notable exception is biology and other life sciences, where women have a higher likelihood of earning a bachelor's degree than men.

The second panel of Table 9 presents estimates of STEM major completion with pre-college controls. Holding constant family income and prior test scores, the magnitude of the gender differences is reduced by one-half or more, though the general pattern still holds.

The finding that women are less likely to earn bachelor's degrees in math, physical sciences and engineering than their male counterparts with equivalent resources and math skills begs the question of whether changes in faculty gender composition would likely alter the outcome. In order to gauge whether student-teacher gender matching affects successful STEM major completion, we estimate models of degree completion which include controls for both middle/high school and first-year-in-college matching of students and instructors, conditional on first-year coursework in college.

Estimates of the determinants of major completion are presented in Table 10. Not surprisingly, first year coursework is strongly related to eventual degree attainment. The greater the number of engineering courses taken in the first year, the more likely a student will eventually earn an engineering degree. The same is true for math and science. Controlling for pre-college influences and course selection in the first year in college, women are no less likely than men to complete a degree in a STEM field, including actually being more likely to complete physical sciences and engineering degrees. Further, while student-teacher gender matching in middle and high school math and science courses is associated with a higher likelihood of first year STEM course taking for women, the same is not true for completion of a STEM major conditional on first year coursework. In fact, the correlation between gender matching of students and teachers in middle and high school and STEM major persistence is frequently negative. Further, none of the student-professor matching variables is positively correlated with the likelihood of completing any degree in STEM or completing a particular STEM major.

D. Decomposition of Completing a College Degree in a STEM Major

In order to gauge the relative importance of the factors that contribute to gender differences in eventual college major completion, we conduct a decomposition analysis in the spirit of Arcidiacono and Koedel (2014). The overall gender gap is the difference in the predicted probabilities of completing a given major for men and for women. These unconditional probabilities are the products of each gender's conditional probabilities at each stage: graduating with a STEM major (conditional on first-year course taking, college entry and all pre-college outcomes), first-year coursework (conditional on entering college and pre-college outcomes), college entry (conditional on earning a high school diploma and pre-high-school test scores) and high school graduation (conditional on test scores in 5^{th} and 8^{th} grade).

The portion of the overall gap that is attributable to gender differences at each stage can be assessed by assigning women the relevant values for men (while keeping the values for women at prior stages constant) and recalculating the differences in predicted STEM major attainment probabilities between men and women. Details are provided in the Appendix.

Results from the decomposition analyses are presented in Table 11. While the absolute differences in the predicted probability of STEM major completion may seem small, it is important to recognize that even for males, the predicted likelihood of graduating from high school, attending college and completing a STEM major is less than five percent. Conditional on college entry, altering women's first-year college course taking patterns to equal those of men only explain only a modest proportion of the overall gender STEM degree gap. The proportion of the gap explained by gender differences is less than one percent overall, but initial college course taking explains a much higher proportion of the gaps in chemistry and physics, nine and fifteen percent, respectively. This difference across subjects is consistent with there being greater numbers of required courses and more strict course sequencing requirements in the physical sciences compared to life sciences and mathematics. Conditional on high school graduation, changing female college attendance rates (in addition to first-year college course-taking choices) only explains an additional four percent of the STEM degree gender gap. Differences in high school graduate rates explain an even smaller proportion of the gap; about one percent. Assigning women the 5^{th} and 8^{th}-grade test scores of men (in addition to equalizing high school graduation, college attendance and first-year college course taking) explains nearly 60 percent of the overall STEM degree completion gender gap. The proportion of the gap explained by pre-high-school test scores is even higher in chemistry, at nearly 68 percent, but it is lower for physics and biology. However, there is still a substantial proportion of the gender gap in major completion that is unexplained by student achievement prior to high school, high school completion and college entry or by the choice of courses during their first year in college.

V. Summary and Conclusions

Growing concern about the low production of college graduates in STEM fields, particularly among minorities and women, has led to a rapidly growing research literature seeking to understand the causes of these disparities and hence provide guidance as to appropriate policies. The focus of this research has been on the experiences of students once they attend college, including the identity of their instructors and the institutions they enroll in. This college-level focus forecloses the possibility that pre-college experiences, such as the quality and identity of middle and high school instructors and peers can shape future major choices in college.

In this paper, we present new data tracking individual students from elementary school through college. Gender gaps in math achievement are generally modest throughout elementary, middle and high school and women are more likely to successfully complete high school and attend college. Once they get to college, however, they are much less likely than males to obtain a bachelor's degree in a STEM field. Although female college students are more likely to complete a major in biology or other life-science fields, they are much less likely than men to earn a degree in engineering or the physical sciences. Exposure to female math and science teachers in middle and high school is correlated with increases in the number of STEM courses taken by female college freshmen. Likewise, students whose middle and high school math and science teachers held degrees in the relevant field, rather than in education, were more likely to take STEM courses as college freshman. However, the gender matching of students and teachers in secondary school is not associated with greater persistence of women in STEM fields after their freshman year. Similarly, conditional on first-year coursework, exposure to female instructors in STEM courses taken during the first year in college is not associated with a greater likelihood of successfully completing a major in a STEM field.

These findings have several important implications for policy and for future research. First, it is important to realize that underrepresentation of women is not uniform across STEM fields; while women constitute a disproportionally low share of engineering, math and physical science graduates, they have a higher likelihood than men of obtaining a degree in the biological sciences (conditional on attending college). Thus, if the goal is to reduce disparities in female representation in STEM fields, focusing policy on math, engineering and the physical sciences is warranted. Second, the results suggest that altering the gender composition of college faculty is unlikely to substantially change the relative numbers of women in STEM fields. Rather, interventions at the middle and high school level appear more likely to influence young women's college course selection and ultimately completion of a STEM major. While manipulating the gender mix of students in secondary math and science classrooms does not appear to increase STEM course taking by women when they enter college, the characteristics of their middle and high school math and science teachers do appear to influence course selection in college. In particular, increasing the proportion of female math and science teachers in high school and hiring more high school math and science teachers with degrees in the relevant subject area are associated with higher rates of STEM course-taking by women during their first year in college.[9] The increased first-year coursework in turn is correlated with the probability of women successfully completing a STEM major.

[9] Whether such a policy would be desirable for women as a whole is another matter. Teachers tend to be paid less than workers in other occupations requiring a college degree, suggesting that encouraging women to become math and science teachers could lower their lifetime earnings.

References

Anelli, Massimo and Giovanni Peri (2019). The Effects of High School Peers' Gender on College Major, College Performance and Income. *The Economic Journal*, 129, 553-602.

Antecol, Heather, Ozkan Eren, and Serkan Ozbeklik (2015). The Effect of Teacher Gender on Student Achievement in Primary School: Evidence from a Randomized Experiment. *Journal of Labor Economics*, 33, 38.

Arcidiacono, Peter and Cory Koedel (2014). Race and College Success: Evidence from Missouri. *American Economic Journal: Applied Economics*, 6(3): 20-57.

Bettinger, Eric (2010). To Be or Not to Be: Major Choices in Budding Scientists. *American Universities in a Global Market*. Ed. Charles Clotfelter. Chicago: University of Chicago Press, pp. 69-98.

Bettinger, Eric and Bridget Terry Long (2005). Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students. *American Economic Review*, 95(2), 152-157.

Bottia, Martha Cecilia, Elizabeth Sterns, Roslyn Arlin Mickelson, Stephanie Moller and Lauren Valentino (2015). "Growing the Roots of STEM Majors: Female Math and Science High School Faculty and the Participation of Students in STEM," *Economics of Education Review*, 45, 14-27.

Carrell, Scott E., Marianne E. Page and James E. West (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap. *The Quarterly Journal of Economics*, 125(3), 1101-1144.

Card, David and A. Abigail Payne (2017). High School Choices and the Gender Gap in STEM. Cambridge, MA: NBER Working Paper No. 23769.

Dee, Thomas S. (2007). Teachers and the Gender Gaps in Student Achievement. *Journal of Human Resources*, 42(3), 528-554.

Dennehy, Tara and Nilanjana Dasgupta (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences*, 114, 5964–5969.

Dickson, Lisa (2010). Race and gender differences in college major choice. *The Annals of the American Academy of Political and Social Science,* 627 (1), 108-124.

Gershenson, Seth, Stephen Holt, and Nicholas Papageorge (2016). Who Believes in Me? The Effect of Student-Teacher Demographic Match on Teacher Expectations. *Economics of Education Review*, 52, 209–24.

Gershenson, Seth, Cassandra Hart, Joshua Hyman, Constance Lindsay, and Nicholas W. Papageorge (2018). *The long-run impacts of same-race teachers*. Cambridge, MA: NBER Working Paper No. 25254.

Greene, William (2004). The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects. *Econometrics Journal*, 7, 98-119.

Fischer, Stefanie (2017). The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence. *Labour Economics*, 46, 211-226.

Fryer, Roland and Steven Levitt (2010). An Empirical Analysis of the Gender Gap in Mathematics. *American Economic Journal: Applied Economics*, 2, 210–40.

Griffith, Amanda L. (2010). Persistence of Women and Minorities in STEM Field Majors: Is it the School that Matters? *Economics of Education Review*, 29(6), 911-922.

Hoffmann, Florian and Philip Oreopoulos (2009). A Professor Like Me: The Influence of Instructor Gender on College Achievement. *Journal of Human Resources*, 44(2), 479-494.

Kokkelenberg, Edward C., and Eshna Sinha (2010). Who Succeeds in STEM Studies? An Analysis of Binghamton University Undergraduate students. *Economics of Education Review*, 29(6), 935-946.

Mansour, Hani, Daniel Rees, Bryson Rintala, and Nathan Wozny (2018). The Effects of Professor Gender on the Post-Graduation Outcomes of Female Students. *IZA Discussion Paper Series*, 11820, 1–59.

Park, Hyunjoon, Jere R. Behrman, and Jaesung Choi (2018). Do Single-Sex Schools Enhance Students' STEM (Science, Technology, Engineering, and Mathematics) Outcomes? *Economics of Education Review*, 62, 35-47.

Price, Joshua (2010). The Effect of Instructor Race and Gender on Student Persistence in STEM Fields. *Economics of Education Review*, 29(6), 901-910.

Sansone, Dario (2017). Why does teacher gender matter?, *Economics of Education Review*, 61, 9–18.

Sansone, Dario (2019): Teacher Characteristics, Student Beliefs and the Gender Gap in STEM Fields. *Educational Evaluation and Policy Analysis,* forthcoming.

Sass, Tim R. (2017). Understanding Racial and Ethnic Differences in STEM Attainment, unpublished manuscript.

Shi, Ying (2018). The Puzzle of Missing Female Engineers: Academic Preparation, Ability Beliefs, and Preferences. *Economics of Education Review*, 64, 129-143.

Speer, Jamin D. (2017). The gender gap in college major: Revisiting the role of pre-college factors. *Labour Economics,* 44, 69–88.

Turner, Sarah and William Bowen (1999). Choice of major: The changing (unchanging) gender gap. *ILR Review,* 52 (2), 289-313.

Winters, Marcus A., Robert C. Haight, Thomas T. Swaim, Katarzyna A. Pickering (2013). The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from Panel Data. *Economics of Education Review*, 34, 69–75.

Table 1 – Descriptive Statistics by 5th-Grade Cohort
(Students with a test score in grade 5 who are enrolled three or more years later in a public school in Grade 9)

	Cohort			
	1997	**1998**	**1999**	**2000**
Female	0.5000	0.4971	0.4954	0.4982
White	0.5292	0.5366	0.5211	0.5119
Black	0.2469	0.2490	0.2483	0.2409
Hispanic	0.1972	0.1847	0.1996	0.2138
Asian	0.0196	0.0215	0.0214	0.0216
Race-other	0.0072	0.0082	0.0097	0.0118
Free Lunch	0.4253	0.4125	0.4146	0.4082
Reduced-Price Lunch	0.0931	0.1036	0.1048	0.1093
5th-grade SSS Normed Math Score	0.0050	0.0075	0.0151	0.0199
5th-grade NRT Normed Math Score			0.0156	0.0221
6th-grade SSS Normed Math Score	-0.5011	-0.5446	0.0661	0.0809
6th-grade NRT Normed Math Score	-0.5981	0.0721	0.0638	0.0779
7th-grade SSS Normed Math Score	-0.5654	0.0850	0.0735	0.0821
7th-grade NRT Normed Math Score	0.0590	0.0736	0.0695	0.0790
8th-grade SSS Normed Math Score	0.0636	0.0801	0.0688	0.0782
8th-grade NRT Normed Math Score	0.0515	0.0703	0.0564	0.0675
9th-grade SSS Normed Math Score	0.1327	0.1269	0.1233	0.1156
9th-grade NRT Normed Math Score	0.1199	0.1081	0.1045	0.0960
10th-grade SSS Normed Math Score	0.0914	0.2122	0.1828	0.2005
10th-grade NRT Normed Math Score	0.0566	0.0637	0.0551	0.0738
Earned Regular HS Diploma within 4 years of Entering Grade 9	0.6299	0.6095	0.6213	0.6429
Earned Regular HS Diploma within 5 years of Entering Grade 9	0.6323	0.6121	0.6247	0.6465

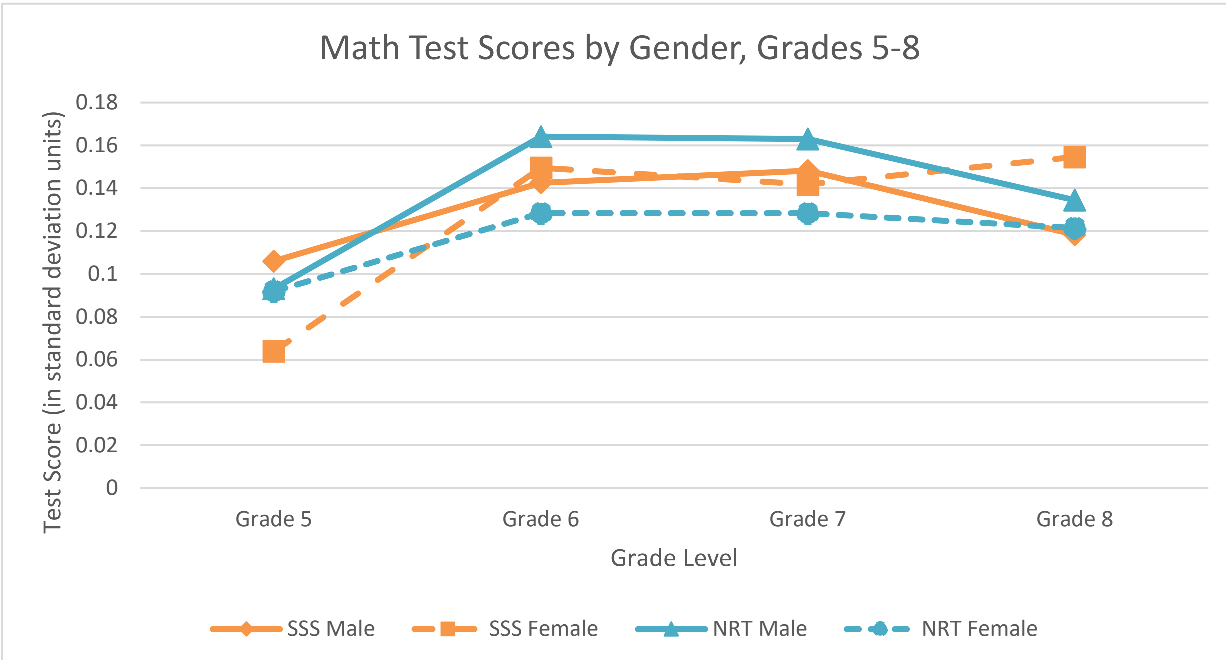
Note: SSS refers to the "Sunshine State Standards" criterion-referenced exam and NRT is the norm-referenced test.

Table 2 - Descriptive Statistics by 5th-Grade Cohort
(Students with a test score in grade 5 who are enrolled exactly four years later in a public school in Grade 9)

	Cohort			
	1997	**1998**	**1999**	**2000**
Female	0.5167	0.5127	0.5111	0.5132
White	0.5408	0.5481	0.5334	0.5249
Black	0.2327	0.2354	0.2344	0.2267
Hispanic	0.1985	0.1856	0.1999	0.2138
Asian	0.0211	0.023	0.0229	0.0231
Race-other	0.0069	0.008	0.0095	0.0114
Free Lunch	0.3985	0.3883	0.3896	0.384
Reduced-Price Lunch	0.0927	0.1035	0.1048	0.1091
5th-grade SSS Normed Math Score	0.0881	0.0797	0.0851	0.0843
5th-grade NRT Normed Math Score			0.0857	0.0925
6th-grade SSS Normed Math Score	-0.2911	-0.5562	0.1345	0.1462
6th-grade NRT Normed Math Score	-0.6128	0.1426	0.133	0.1457
7th-grade SSS Normed Math Score	-0.7875	0.1486	0.1361	0.1449
7th-grade NRT Normed Math Score	0.1309	0.1401	0.1354	0.145
8th-grade SSS Normed Math Score	0.1256	0.1375	0.1296	0.1371
8th-grade NRT Normed Math Score	0.1165	0.1309	0.1185	0.1277
9th-grade SSS Normed Math Score	0.1933	0.1797	0.1768	0.1682
9th-grade NRT Normed Math Score	0.1798	0.1628	0.1581	0.1501
10th-grade SSS Normed Math Score	0.1232	0.2457	0.2143	0.2327
10th-grade NRT Normed Math Score	0.0916	0.0956	0.0884	0.1077
Earned Regular HS Diploma within 4 years of Entering Grade 9	0.6799	0.6537	0.6652	0.6833
Earned Regular HS Diploma within 5 years of Entering Grade 9	0.6823	0.6563	0.6686	0.6867

Note: SSS refers to the "Sunshine State Standards" criterion-referenced exam and NRT is the norm-referenced test.

Figure 1 - Normed Test Scores in Grades 5-8 by Gender
(Students with a test score in grade 5 in 2000 who are enrolled four years later in a public school in Grade 9)



Note: SSS refers to the "Sunshine State Standards" criterion-referenced exam and NRT is the norm-referenced test.

Table 3 – High School Exit by Gender
(Students with a test score in grade 5 in 1997-2000 who are enrolled four years later in a public school in Grade 9)

Gender	No Exit (Enrolled in Each of Grades 9-12)	Dropped out	Exit to home school	Exit to private school	Exit – other
Male	162,205 [64.87]	20,590 [8.23]	2,488 [1.00]	5,380 [2.15]	59,381 [23.75]
Female	184,154 [69.87]	18,038 [6.84]	3,323 [1.26]	5,226 [1.98]	52,832 [20.04]

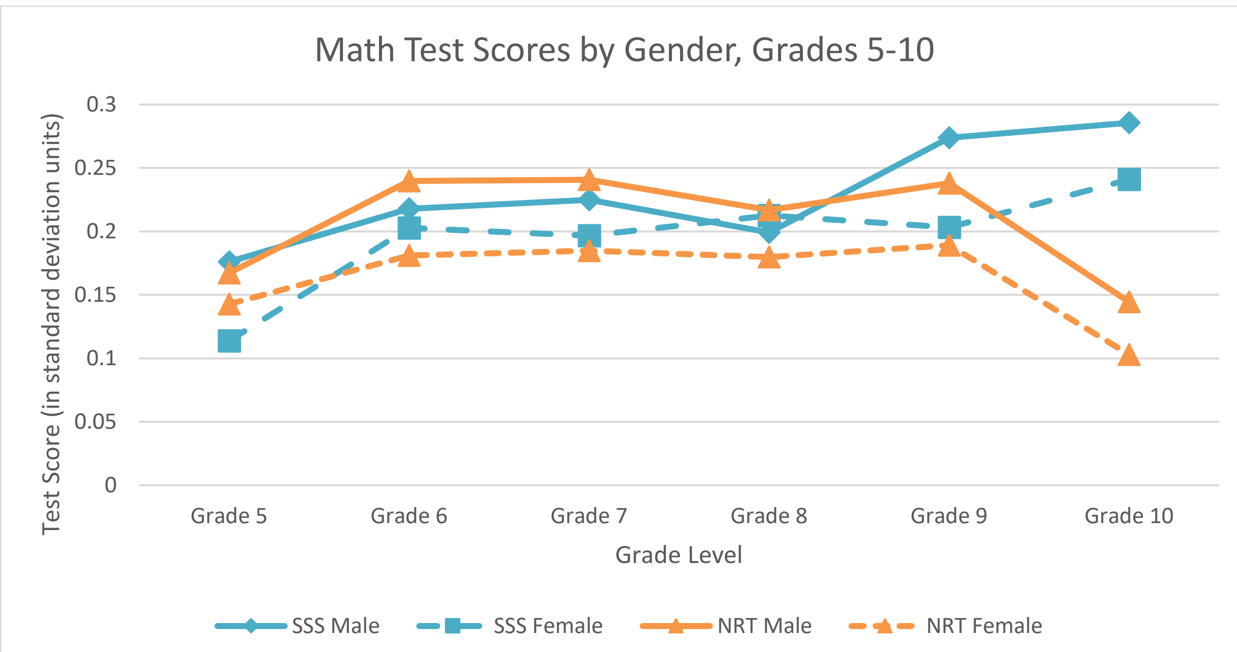
Note: numbers in brackets are row percentages.

Table 4 – Regular High School Diploma Receipt within 5 Years of Entering Grade 9 by Gender
(Students with a test score in grade 5 in 1997-2000 who are enrolled four years later in a public school in Grade 9)

Gender	Did Not Receive Diploma	Received Diploma
Male	89,301 [35.71]	160,743 [64.29]
Female	78,625 [29.83]	184,948 [70.17]

Note: numbers in brackets are row percentages.

Figure 2 – Normed Test Scores in Grades 5-10 by Gender
(Students with a test score in grade 5 in 2000 who are enrolled four years later in a public school in Grade 9 and continue to be enrolled in Grades 10 and 11)



Note: SSS refers to the "Sunshine State Standards" criterion-referenced exam and NRT is the norm-referenced test.

Table 5 – College Attendance in Year Immediately Following Receipt of Regular High School Diploma by Race/Ethnicity and Gender
(Students with a test score in grade 5 in 1997 who are enrolled four years later in a public school in Grade 9 and graduate within 5 years with a regular high school diploma)

Race/Ethnicity and Gender	No College	FL Community College	4-year FL Public University	4-year FL Private College/Univ	4-year College Out of State
Male	10,610 [35.72]	8,431 [28.38]	6,519 [21.95]	934 [3.14]	934 [3.14]
Female	9,802 [28.02]	11,176 [31.95]	9,402 [26.88]	1,139 [3.26]	3,460 [9.89]

Note: numbers in brackets are row percentages.

Table 6 – Probit Estimates of the Determinants of Taking One or More Courses in a STEM Field in the First Year in College
(Students with a test score in grade 5 in 1997-2000 who are enrolled four years later in a public school in Grade 9 and graduate within 5 years with a regular high school diploma and attend a Florida public university within one year of receiving their diploma)

Explanatory Variables	Estimated Marginal Effect	
Female	-0.0359**	-0.0325**
	(0.0028)	(0.0030)
Black	0.0100*	0.0227**
	(0.0038)	(0.0045)
Hispanic	0.0530**	0.0588**
	(0.0034)	(0.0037)
Asian	0.0304**	0.0270**
	(0.0058)	(0.0061)
Other Race	0.0142	0.0089
	(0.0183)	(0.0199)
Free Lunch		-0.0084
		(0.0046)
Reduced-Price Lunch		-0.0103
		(0.0061)
SSS Grade 9		0.0065
		(0.0044)
NRT Grade 9		0.0019
		(0.0032)
SSS Grade 10		0.0015
		(0.0056)
NRT Grade 10		0.0029
		(0.0029)
NRT Grade 7		0.0015
		(0.0032)
SSS Grade 8		0.0150**
		(0.0051)
NRT Grade 8		-0.0031
		(0.0034)
SSS Grade 5		-0.0134**
		(0.0036)
Observations	74,528	67,297

Note: SSS refers to the "Sunshine State Standards" criterion-referenced exam and NRT is the norm-referenced test. Excludes remedial math courses. All models include cohort controls. Reported estimates are marginal effects. Standard errors in parentheses. * significant at the 5% level, ** significant at the 1% level in a two-tailed test.

Table 7 – Probit Estimates of Gender Differences in the Probability of Taking One or More Courses in a STEM Field in the First Year in College
(Students with a test score in grade 5 in 1997-2000 who are enrolled four years later in a public school in Grade 9 and graduate high school within 5 years and attend a Florida public university within one year of receiving their diploma)

Subject	Estimated Female Marginal Effect
Any STEM	-0.0325** (0.0030)
Biology	0.0613** (0.0034)
Chemistry	-0.0048 (0.0034)
Computer Science	-0.0238** (0.0011)
Engineering	-0.0820** (0.0020)
Math	-0.0514** (0.0037)
Physics	-0.0313** (0.0014)
Statistics	0.0512** (0.0027)
Observations	67,297

Note: SSS refers to the "Sunshine State Standards" criterion-referenced exam and NRT is the norm-referenced test. Excludes remedial math courses. All models include controls for race/ethnicity, lunch status, math test scores in grades 5-10 and cohort controls (as in the second column of Table 6). Standard errors in parentheses. *significant at the 5% level, **significant at the 1% level in a two-tailed test.

Table 8 – Probit Estimates of the Determinants of Taking >=1 Courses in any STEM Field in the First Year in College (Students with a test score in grade 5 in 1997-2000 who are enrolled four years later in a public school in Grade 9 and graduate high school within 5 years and attend a Florida public university within one year of receiving their diploma)

Explanatory Variables	Probit Without HS Fixed Effects	Linear Probability Without HS Fixed Effects	Linear Probability With HS Fixed Effects	Linear Probability with HS Fixed Effects (Schools with 100+ Teachers)	Probit IV
Female	0.0093	0.0105	-0.0330	-0.0799**	-0.6874**
	(0.0246)	(0.0251)	(0.0255)	(0.0300)	(0.1591)
Female x Proportion of Middle and HS Math and Science Courses Taught by a Female Teacher	0.0607**	0.0623**	-0.0005	-0.0054	1.4025**
	(0.0145)	(0.0148)	(0.0151)	(0.0175)	(0.2179)
Enrolled in at Least One HS Biology Course Taught by a Teacher with a BA in Biology	0.0142*	0.0147**	-0.0152*	-0.0107	-0.1478
	(0.0060)	(0.0062)	(0.0067)	(0.0076)	(0.0793)
Enrolled in at Least One HS Chemistry Course Taught by a Teacher with a BA in Chemistry	0.0203*	0.0206**	0.0142	0.0223*	0.0214
	(0.0075)	(0.0078)	(0.0086)	(0.0093)	(0.0739)
Enrolled in at Least One HS Physics Course Taught by a Teacher with a BA in Physics	-0.0131	-0.0137	-0.0163	-0.0168	0.3983**
	(0.0124)	(0.0118)	(0.0132)	(0.0133)	(0.1343)
Enrolled in at Least One HS Math Course Taught by a Teacher with a BA in Math	0.0132*	0.0134*	0.0069	0.0016	0.1538**
	(0.0051)	(0.0053)	(0.0058)	(0.0065)	(0.0526)
Female x Proportion of Female Students in Middle and HS Math and Science Courses	-0.1467**	-0.1502**	-0.0045	0.0924	-0.5112**
	(0.0405)	(0.0416)	(0.0427)	(0.0503)	(0.1600)
Observations	49,633	49,633	49,627	33,230	49,570

Note: Excludes remedial math courses. All models include cohort controls, student race/ethnicity indicators, student race/ethnicity interactions with teacher race/ethnicity, student race/ethnicity interactions with peer student race/ethnicity, controls for free and reduced-price lunch in grade 5 and controls for math test scores in grades 5, 7, 8. Reported estimates are marginal effects. Standard errors in parentheses. *significant at the 5% level, ** significant at the 1% level in a two-tailed test.

Table 9 – Probit Estimates of the Determinants of Earning a Bachelor's Degree in a STEM Major Within 9 years of Starting Grade 9 – *With and Without Pre-College Controls*
(Students with a test score in grade 5 in 1997 who are enrolled four years later in a public school in Grade 9 and graduate within 5 years with a regular high school diploma and attend a Florida public university within one year of receiving their diploma)

Explanatory Variables	Earning a Bachelor's Degree in STEM	Earning a Bachelor's Degree in Bio. Sci.	Earning a Bachelor's Degree in Chemistry	Earning a Bachelor's Degree in Engineering	Earning a Bachelor's Degree in Math	Earning a Bachelor's Degree in Physics
No Pre-College Controls						
Female	-0.0475**	0.0016**	-0.0022**	-0.0412**	-0.0022**	-0.0030**
	(0.0019)	(0.0004)	(0.0005)	(0.0013)	(0.0004)	(0.0004)
Observations	75,292	75,292	75,292	75,292	75,292	62,816
With Pre-College Controls for Family Income and Student Achievement						
Female	-0.0254**	0.0020**	-0.0004	-0.0227**	-0.0005*	-0.0011**
	(0.0017)	(0.0005)	(0.0004)	(0.0011)	(0.0002)	(0.0002)
Observations	67,966	67,966	67,966	67,966	67,966	56,776

Note: All models include cohort controls and controls for race/ethnicity. Estimates in the lower panel are from models that include controls for free and reduced-price lunch in grade 5 and controls for math test scores in grades 5, 7, 8, 9 and 10. Reported estimates are marginal effects. Standard errors are in parentheses. *significant at the 5% level, ** significant at the 1% level in a two-tailed test.

Table 10 – Probit Estimates of the Determinants of Earning a Bachelor's Degree in a STEM Major Within 9 years of Starting Grade 9 (Students with a test score in grade 5 in 1997 who are enrolled four years later in a public school in Grade 9 and graduate high school within 5 years and attend a Florida public university within one year of receiving their diploma)

Explanatory Variables	Bachelor's Degree in any STEM	Bachelor's Degree in Bio. Sci.	Bachelor's Degree in Chemistry	Bachelor's Degree in Engineering	Bachelor's Degree in Math	Bachelor's Degree in Physics
Female	0.0183 (0.0132)	0.0020 (0.0030)	0.0018 (0.0026)	0.0189** (0.0071)	-0.0003 (0.0018)	0.0100** (0.0099)
Female x Proportion of Middle and HS Math and Science Courses Taught by a Female Teacher	-0.0186* (0.0082)	0.0004 (0.0019)	-0.0036* (0.0016)	-0.0072 (0.0039)	-0.0001 (0.0010)	-0.0003 (0.0005)
Enrolled in at Least One HS Biology Course Taught by a Teacher with a BA in Biology	-0.0054 (0.0029)	0.0006 (0.0009)	-0.0003 (0.0006)	-0.0022 (0.0010)	-0.0004 (0.0003)	0.0001 (0.0002)
Enrolled in at Least One HS Chemistry Course Taught by a Teacher with a BA in Chemistry	0.0089* (0.0042)	-0.0002 (0.0010)	-0.0009 (0.0005)	-0.0006 (0.0014)	0.0005 (0.0006)	0.0001 (0.0002)
Enrolled in at Least One HS Physics Course Taught by a Teacher with a BA in Physics	0.0152** (0.0064)	-0.0025 (0.0007)	0.0015 (0.0014)	0.0063** (0.0028)	0.0024* (0.0014)	0.0005 (0.0005)
Enrolled in at Least One HS Math Course Taught by a Teacher with a BA in Math	0.0011 (0.0026)	-0.0008 (0.0006)	0.0009 (0.0006)	-0.0004 (0.0009)	0.0000 (0.0003)	0.0000 (0.0001)
Female x Proportion of Female Students in Middle and HS Math and Science Courses	-0.0367 (0.0230)	-0.0004 (0.0052)	-0.0007 (0.0045)	-0.0459** (0.0113)	0.0004 (0.0028)	-0.0045** (0.0016)
Female x Proportion of First-Year College STEM Courses Taught by a Female Instructor	-0.0156** (0.0042)	-0.0029** (0.0010)	-0.0014 (0.0009)	-0.0059** (0.0021)	-0.0009 (0.0006)	0.0001 (0.0002)
No. of Computer Courses in First Year	0.0228** (0.0022)	-0.0024 (0.0023)	-0.0016 (0.0012)	0.0001 (0.0009)	0.0004** (0.0002)	0.0000 (0.0001)
No. of Engineering Courses in First Year	0.0187** (0.0010)	-0.0011 (0.0005)	-0.0011** (0.0004)	0.0076** (0.0005)	-0.0004* (0.0002)	-0.0001* (0.0000)
No. of Math Courses in First Year	0.0052** (0.0010)	-0.0000 (0.0003)	0.0000 (0.0002)	0.0027** (0.0004)	0.0006** (0.0001)	0.0001 (0.0000)
No. of Statistics Courses in First Year	-0.0277** (0.0025)	-0.0012* (0.0006)	-0.0019** (0.0005)	-0.0129** (0.0013)	0.0001 (0.0002)	-0.0005** (0.0002)
No. of Science Courses in First Year	0.0113** (0.0005)	0.0010** (0.0001)	0.0010** (0.0001)	0.0015** (0.0002)	0.0000 (0.0001)	0.0001** (0.0000)

Observations	48,550	48,281	48,550	48,550	48,550	39,867

Note: All models include cohort controls, controls for free and reduced-price lunch in grade 5 and controls for math test scores in grades 5, 7, and 8. Reported estimates are marginal effects. Standard errors are in parentheses. *significant at the 5% level, ** significant at the 1% level in a two-tailed test. Excludes remedial math courses. All models include cohort controls.

Table 11 – Decomposition of Male-Female STEM Degree Completion Gap

Degree	Predicted Degree Completion Probability for Males (Percent)	Predicted Degree Completion Probability for Females (Percent)	Male-Female Gap in Predicted Degree Completion Probability (Percentage Points)	Percentage of Predicted Gap Explained by:				
				Pre-High School Math Test Scores	High School Graduation	College Entry	First-Year College Course-work	Other Factors
Any STEM	.04672	.02844	0.01828	59.72	1.05	3.75	0.79	34.68
Chemistry	.00076	.00050	0.00026	67.81	1.20	4.26	9.15	17.59
Physics	.00005	.00000	0.00004	25.31	0.45	1.59	15.21	57.45
Biology	.00052	.00021	0.00031	38.87	0.69	2.44	-4.33	62.33

Appendix

Following Arcidiacono and Koedel (AK), the de-composition involves a multi-step process. We begin with an analog of AK's equation (1), which represents the probability of graduating from college with a STEM major:

$$\sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} Pr(Y = 1|f, c, d, x, g) \times Pr(f|c, d, x, g) \times Pr(c|d, x, g) \times Pr(d|x, g) \times \Pr(x|g) \qquad [1]$$

where

Y= graduate from college with a STEM major
f = first-year coursework
c =enroll in college
d = student earns a high school diploma
x = pre-HS academic background (measured by test scores)
g =gender

The overall difference in graduation rates between males and females (D_g) therefore equals:

$$D_g = \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, male) \times \Pr(f|c, d, x, male) \times \Pr(c|d, x, male) \times \Pr(d|x, male) \times \Pr(x|male) -$$
$$\sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, female) \times \Pr(f|c, d, x, female) \times \Pr(c|d, x, female) \times \Pr(d|x, female) \times \Pr(x|female)$$

The first step in the decomposition is to determine how that difference in predicted probabilities would change if women had the same conditional stem graduation rate as men, but their own true values for all of the other components (f, c, d, x):

$$D_{gd} = \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, male) \times \Pr(f|c, d, x, female) \times \Pr(\mathrm{c}|d, x, female) \times \Pr(d|x, female) \times \Pr(x|female) - \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, female) \times \Pr(f|c, d, x, female) \times \Pr(\mathrm{c}|d, x, female) \times \Pr(d|x, female) \times \Pr(x|female)$$

The next step in the decomposition is to determine how that difference in predicted probabilities would change if women also had the same first-year college coursework as men, but their own true values for all of the other components (c, d, x):

$$D_{fd} = \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, male) \times \Pr(f|c, d, x, male) \times \Pr(\mathrm{c}|d, x, female) \times \Pr(d|x, female) \times \Pr(x|female) - \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, male) \times \Pr(f|c, d, x, female) \times \Pr(\mathrm{c}|d, x, female) \times \Pr(d|x, female) \times \Pr(x|female)$$

The third step is to determine how the gap would change further if women had both the same first-year coursework as men and the same probability of attending college:

$$D_{cd} = \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, male) \times \Pr(f|c, d, x, male) \times \Pr(\mathrm{c}|d, x, male) \times \Pr(d|x, female) \times \Pr(x|female) - \sum_{x \in X} \sum_{d \in D} \sum_{c \in C} \sum_{f \in F} \Pr(Y = 1|f, c, d, x, male) \times \Pr(f|c, d, x, male) \times \Pr(\mathrm{c}|d, x, female) \times \Pr(d|x, female) \times \Pr(x|female)$$

The effects of earning a high school diploma and pre-high-school achievement are determining by continuing in a similar recursive fashion.