Making Connections

November 2017

Associations between predictive indicators and postsecondary science, technology, engineering, and math success among Hispanic students in Texas

Trisha Borman
Jonathan Margolin
American Institutes for Research

Marshall Garland
Amie Rapaport
Gibson Consulting Group

So Jung Park
Christina LiCalsi
American Institutes for Research

Key findings

- For Texas public high school graduates enrolled in a two-year or a four-year college, indicators of academic experiences, achievement in math and science, and high school attendance rate were strongly associated with postsecondary science, technology, engineering, and math (STEM) success. The associations were generally similar for Hispanic students and non-Hispanic White students.
- The indicators associated with postsecondary STEM success included number of math or science courses taken, number of Advanced Placement math or science courses taken, highest math or science course taken, and scores on state assessments.

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL SOUTHWEST
Regional Educational Laboratory
At SEDL

U.S. Department of Education
Betsy DeVos, *Secretary*

Institute of Education Sciences
Thomas W. Brock, *Commissioner for Education Research*
Delegated the Duties of Director

National Center for Education Evaluation and Regional Assistance
Ricky Takai, *Acting Commissioner*
Elizabeth Eisner, *Acting Associate Commissioner*
Amy Johnson, *Action Editor*
Chris Boccanfuso, *Project Officer*

REL 2018–279

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

November 2017

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0012 by Regional Educational Laboratory Southwest administered by SEDL. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Borman, T., Margolin, J., Garland, M., Rapaport, A., Park, S. J., & LiCalsi, C. (2017). *Associations between predictive indicators and postsecondary science, technology, engineering, and math success among Hispanic students in Texas* (REL 2018–279). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

Summary

Nationwide, Hispanic students continue to be underrepresented among students who complete a four-year degree in science, technology, engineering, and math (STEM) fields (National Center for Education Statistics, 2016) and among workers in STEM fields (Beede et al., 2011). This discrepancy is a concern, especially in light of the projected growth in employment in STEM fields (about a million jobs from 2012 to 2022; Vilorio, 2014) and in light of the fact that wages for jobs in STEM fields are 26 percent higher on average than wages for jobs in non-STEM fields (Langdon, McKittrick, Beede, Khan, & Doms 2011). Concern is particularly acute in Texas, where Hispanic students account for 51 percent of the K–12 student population (Texas Education Agency, 2016).

The Texas Hispanic STEM Research Alliance, made up of representatives from Regional Education Service Centers, school districts, postsecondary institutions, and state education agencies, partnered with the Regional Educational Laboratory Southwest on a series of studies of predictors of postsecondary STEM success (defined as enrolling in, persisting in, or completing a postsecondary STEM major or degree) for Hispanic students in Texas. The first study reviewed the research literature to identify indicators (malleable factors that can be measured in K–12 settings) that predict students' postsecondary STEM success (Hinojosa, Rapaport, Jaciw, LiCalsi, & Zacamy, 2016). Among the indicators identified were the number of math courses taken, the number of science courses taken, the level of math or science courses taken, interest or confidence in STEM, achievement in middle school and high school (for example, grade point average and scores on standardized math and science assessment tests), and SAT and ACT math scores. This report presents the findings of the second study, which examined the associations between some of the indicators identified in the first study and postsecondary STEM success among Texas students.

The current study used data on seven cohorts of students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college or a four-year Texas public or private college or university by spring 2011. Key findings include:

- Math and science courses taken in high school showed statistically significant and robust associations with postsecondary STEM success for students who enrolled in a two-year or a four-year college.
  - Among students who enrolled in a two-year college, taking higher-level math or science courses (higher level than the typical student takes) was associated with a greater likelihood of declaring and persisting in a STEM major and of completing a STEM degree. The converse was also true: taking math courses that were less advanced than the typical student takes was associated with a lower likelihood of declaring a STEM major and completing a STEM degree. Taking more Advanced Placement science courses was associated with a greater likelihood of declaring a STEM major, and taking more science courses was associated with a greater likelihood of completing a STEM degree.
  - Among students who enrolled in a four-year college, taking an Advanced Placement calculus course and taking Advanced Placement Physics were associated with a greater likelihood of persisting in a STEM major. Taking more math courses and taking more science courses were associated with a greater likelihood of persisting in a STEM major, and taking more Advanced Placement math courses was associated with a greater likelihood of completing a STEM degree.

- Achievement in math and science in high school showed statistically significant and robust associations with postsecondary STEM success for students who enrolled in a two-year or a four-year college.
    - Among students who enrolled in a two-year college, higher grade 11 state assessment scores in math and science were associated with a greater likelihood of declaring and persisting in a STEM major and completing a STEM degree.
    - Among students who enrolled in a four-year college, a higher grade 10 state assessment score in science was associated with a greater likelihood of persisting in a STEM major, and a higher grade 11 score was associated with a greater likelihood of completing a STEM degree.
- A higher high school attendance rate was associated with a greater likelihood of declaring a STEM major and completing a STEM degree among students who enrolled in a two-year college and with a greater likelihood of persisting in a STEM major among students who enrolled in a four-year college.
- No indicators were predictive only for Hispanic students. However, among students who enrolled in a two-year college, the grade 11 state assessment score in science was less strongly associated with the likelihood of declaring a STEM major for Hispanic students than for non-Hispanic White students.

The finding that indicators that predict postsecondary STEM success function similarly for Hispanic and non-Hispanic students suggests that one explanation for the lower percentage of Hispanic students who declare and persist in a STEM major and complete a STEM degree might be that Hispanic students participate in advanced, rigorous math and science courses in high school at a lower rate than non-Hispanic White students do. For example, among students who enrolled in a four-year college, an Advanced Placement calculus course was the highest math course taken for 20 percent of Hispanic students and 27 percent of non-Hispanic White students, and Physics was the highest science course taken for 49 percent of Hispanic students and 54 percent of non-Hispanic White students (an additional 8 percent of whom took Advanced Placement Physics). This study lends urgency to policies and practices that address underrepresentation of Hispanic students in rigorous courses.

Contents

Tables

Why this study?

Nationwide, Hispanic students continue to be underrepresented among workers in science, technology, engineering, and math (STEM) fields. In 2009 Hispanic workers accounted for 14 percent of the workforce but held only 6 percent of jobs in STEM fields (Beede et al., 2011). This discrepancy is a concern, especially in light of the projected growth in employment in STEM fields (about a million jobs from 2012 to 2022; Vilorio, 2014) and in light of the fact that wages for jobs in STEM fields are 26 percent higher on average than wages for jobs in non-STEM fields (Langdon et al., 2011). Moreover, this wage difference is larger for Hispanic workers than for non-Hispanic White workers (Beede et al., 2011).

The discrepancy in employment is likely to persist if current trends continue. Hispanic students received 11 percent of bachelor's degrees conferred in 2013 but only 9 percent of bachelor's degrees in STEM majors (National Center for Education Statistics, 2016). Among students who declare a STEM major, Hispanic students persist at lower rates than non-Hispanic White students do (Higher Education Research Institute, 2010). Concern is particularly acute in Texas, where Hispanic students account for 51 percent of the K–12 student population (Texas Education Agency, 2016).

Identifying indicators of postsecondary STEM outcomes that differ specifically for Hispanic students could help explain why a lower percentage of Hispanic than of non-Hispanic students declare and persist in a STEM major and complete a STEM degree

The Texas Hispanic STEM Research Alliance, made up of representatives from Regional Education Service Centers, school districts, postsecondary institutions, and state education agencies, partnered with the Regional Educational Laboratory Southwest on a series of studies to identify indicators (malleable factors that can be measured in K–12 settings) that predict postsecondary STEM success (defined as enrolling in, persisting in, or completing a postsecondary STEM major or degree) for Hispanic students in Texas. Identifying leading indicators (defined as "indicators that provide early signals of progress toward academic achievement"; Foley et al., 2008, p. 1) is a first step in developing effective education interventions (Neild, Balfanz, and Herzog, 2007) and may suggest areas for further research.[1]

The first study in the series, a systematic literature review, found that the number of math courses taken, the number of science courses taken, the level of those courses, and students' interest in STEM and confidence in their ability to succeed in STEM courses predict postsecondary STEM success for all student subgroups (Hinojosa et al., 2016). The review found only limited research on K–12 predictors of postsecondary STEM success specifically for Hispanic students (either as a separate group or as part of a larger racial/ethnic minority group), though four studies suggested that the number of math and science courses taken and the grades earned in those courses are less predictive of pursuing a STEM major for racial/ethnic minority students than for non-Hispanic White students. How much, if at all, the relationship between predictors and success varies specifically for Hispanic students was unclear. The review also found that despite having similar interest in STEM and confidence in ability to succeed in STEM courses, racial/ethnic minority students are less likely than non-Hispanic White students to take the highest math and science courses in high school and to enroll and persist in a STEM major and complete a STEM degree.

This study, the second in the series, sought to identify the indicators that predict postsecondary STEM outcomes among Hispanic students in Texas. In light of the fact that nearly all the indicators identified by Hinojosa et al.'s (2016) recent literature review focused on academic experiences in high school, the current study considered indicators of experiences during, but not prior to, high school. Identifying indicators that differ specifically for Hispanic students could help explain why a lower percentage of Hispanic than of non-Hispanic White students declare and persist in a STEM major and complete a STEM degree. For example, if enrolling in rigorous math courses in high school is a strong predicator of postsecondary STEM success, differences in enrollment rates between Hispanic students and non-Hispanic White students should be associated with differences in their postsecondary STEM success. If not, some other factor may account for the differences.

What the study examined

The study addressed two research questions:

- Which indicators of academic experiences in high school predict postsecondary STEM success among Texas high school graduates?
- Do indicators that predict postsecondary STEM success function differently for Hispanic students (of any racial group) than for non-Hispanic White students?

The study team also explored whether indicators that predict postsecondary STEM success function differently for female and male students and before and after a change in state graduation requirements that began in 2004/05. Supplementary findings related to these topics are in appendix A.

The study examined data on seven cohorts of students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college or a four-year Texas public or private college or university by spring 2011. Students were examined separately by type of college they enrolled in (see box 1 for a summary of the data and methodology and appendix B for details). There was a modest but consistent gap between Hispanic students and non-Hispanic White students in the study sample who enrolled in a four-year college with respect to persisting in a STEM major and completing a STEM degree (but not in declaring a STEM major; see box 2 and table C9 in appendix C). There were no gaps among students who enrolled in a two-year college (see table C8 in appendix C). These discrepancies were similar to those that have been identified in other studies (Higher Education Research Institute, 2010; National Center for Education Statistics, 2016).

There was a modest but consistent gap between Hispanic students and non-Hispanic White students in the study sample who enrolled in a four-year college with respect to persisting in a STEM major and completing a STEM degree but not in declaring a STEM major

The study explored the following indicators:

- Number of math and number of science courses taken in high school.
- Highest math and highest science course taken in high school.
- Number of Advanced Placement math and science courses taken in high school.
- SAT composite score (sum of math and verbal subtest scores).
- State assessment scores in math and science in grades 10 and 11.
- High school attendance rate (days attended as a percentage of days enrolled).

The choice of indicators was informed by the findings of the literature review (Hinojosa et al., 2016), though data were not available for the current study for all the indicators identified in the review (such as student interest in STEM and parent encouragement of science or math education). The current study also examined the high school attendance rate, which was not identified in the review but fits the definition of an indicator as a malleable factor associated with measures of academic achievement (Ginsburg, Jordan, & Chang, 2014) and high school graduation (Allensworth & Easton, 2007).

The outcomes of interest were three binary (yes or no) measures of postsecondary success in STEM:

- Declaring a STEM major: whether a college student ever declared a STEM major.
- Persisting in a STEM major: whether a college student who declared a STEM major remained enrolled as a STEM major for all subsequent semesters.
- Completing a STEM degree: whether a college student completed a degree (even from a different institution of higher education than at first enrollment) with a STEM major.

Box 1. Data and methodology

Data

The study team linked high school student academic data, which were provided by the Texas Education Agency, to college enrollment data, which were provided by the Texas Higher Education Coordinating Board, on the seven cohorts of students who entered grade 9 for the first time in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a Texas college between 2004 and spring 2011.[1] Students who entered a Texas public school after grade 9 were folded into the cohort corresponding to their grade–school year configuration, and students who did not progress sequentially through high school (for example, because they repeated or skipped a grade) were not reassigned to a different cohort.

Students were split into two samples by whether they first enrolled in a two-year Texas public college (a two-year college) or a four-year Texas public or private college or university (a four-year college). The sample of students who enrolled in a two-year college included 649,217 students, of whom approximately 46 percent were non-Hispanic White, 38 percent were Hispanic, and 13 percent were non-Hispanic Black. The sample of students who enrolled in a four-year college included 455,560 students, of whom approximately 55 percent were non-Hispanic White, 25 percent were Hispanic, 14 percent were non-Hispanic Black, 83 percent enrolled in a public college or university, and 17 percent enrolled in a private college or university. The analyses of students who enrolled in a two-year college examined whether they completed a two-year or a four-year STEM degree, and the analyses of students who enrolled in a four-year college examined whether they completed a four-year STEM degree (regardless of whether they previously or subsequently completed a two-year STEM degree).

Data on whether students declared a STEM major and on whether students persisted in a STEM major were not available for students who enrolled in a four-year private college. Data on whether students who enrolled in a two-year or a four-year public college persisted in a STEM major and on whether students who enrolled in a two-year college completed a STEM degree were not available for cohort 7 (which entered grade 9 in 2006), and data on whether students who enrolled in a four-year college completed a STEM degree were not available for cohorts 4–7 (which entered grade 9 in 2003–06). See appendix B for a detailed description of the data and sample.

Methodology

Research question 1 was addressed using regression models that examined relationships between predictive indicators and postsecondary STEM outcomes, while nonmalleable school-level contextual variables, student-level covariates (such as race/ethnicity, sex, and English learner status), and cohort fixed effects were controlled for (see table B2 in appendix B). Research question 2 was addressed by adding interaction terms to the regression models. The interaction terms provided a separate estimate of each indicator's association with each outcome for Hispanic students, non-Hispanic Black students, non-Hispanic White students, and students of other races/ethnicities.

Because factors excluded from the regression models might influence results, the associations between indicators and outcomes were tested to consider whether they were robust to changes in covariates used in the models and changes in the samples. A detailed technical description of the analytic methodology is in appendix B.

Note

1. Approximately 54 percent of students who entered grade 9 for the first time in 2000–06 and were enrolled in a Texas public high school for at least three years had enrolled in a Texas college by spring 2011. In 2007/08 approximately 93 percent of Texas high school graduates who enrolled in a college or university during the fall after high school graduation attended a school in Texas (Texas Guaranteed Student Loan Corporation, 2013).

Box 2. Summary of postsecondary outcomes across racial/ethnic groups

Racial/ethnic disparities in postsecondary outcomes identified in previous studies (Higher Education Research Institute, 2010; National Center for Education Statistics, 2016) existed for some of the outcomes among students in the study sample. Findings for non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or other Pacific Islander students, who are included in the "other" category, are not discussed here because these students account for a low proportion of the student population.

Among students who enrolled in a two-year or a four-year college, the rates at which Hispanic students and non-Hispanic White students declared a STEM major were similar and slightly higher than the rates at which non-Hispanic Black students did (see tables C8 and C9 in appendix C).

Among students who enrolled in a two-year college, the rate at which Hispanic students and non-Hispanic White students persisted in a STEM major were similar, but among students who enrolled in a four-year college, Hispanic students persisted in a STEM major (53 percent) at a lower rate than non-Hispanic White students did (56 percent). Among students who enrolled in a two-year college or a four-year college, Hispanic students had higher rates of persisting in a STEM major than non-Hispanic Black students did (see tables C8 and C9 appendix C).

Among students who enrolled in a four-year college, Hispanic students earned a degree in any major at a substantially lower rate (37 percent) than non-Hispanic White students did (57 percent) and at a higher rate than non-Hispanic Black students did (28 percent). The pattern was similar among students who enrolled in a two-year college, although the degree completion rate was much lower for this sample, with correspondingly smaller differences across groups. Among students who enrolled in a four-year college, Hispanic students completed a STEM degree at a lower rate (7 percent) than non-Hispanic White students did (10 percent) but at a higher rate than non-Hispanic Black students did (3 percent). The rates at which students completed a STEM degree were too low to discern clear differences across groups (see tables C8 and C9 in appendix C).

A wide range of indicators in math and science predicted postsecondary STEM success for students who enrolled in a two-year college and for students who enrolled in a four-year college

What the study found

This section presents the key findings about the indicators that predict postsecondary STEM success among a sample of students who attended Texas public high schools and who subsequently enrolled in a two- or four-year college in Texas. It also describes whether the indicators that predict postsecondary STEM success function differently for Hispanic students than for non-Hispanic White students.

Students' academic experiences and achievement in math and science in high school predicted postsecondary STEM success

When analyzed jointly, the full set of indicators of academic experiences and achievement in math and science in high school predicted postsecondary STEM success.[2] Indeed, a wide range of indicators in both subjects predicted postsecondary STEM success for students who enrolled in a two-year college and for students who enrolled in a four-year college (table 1). All results with a checkmark in table 1 are both statistically significant and robust to various model specifications. For comparisons related to the highest math course taken, the reference group was students whose highest math course taken was Algebra II; for comparisons related to the highest science course taken, the reference group

Table 1. Associations between indicators of experiences in math and science in high school and postsecondary STEM success among Texas high school graduates

Indicator	Students who enrolled in a two year college			Students who enrolled in a four year college		
	Declare a STEM major	Persist in a STEM major	Complete a STEM degree	Declare a STEM major	Persist in a STEM major	Complete a STEM degree
Highest math course taken in high school (reference category is Algebra II)						
Algebra I	–✔		–✔	✔		
Mathematics Models						
Geometry			–✔			
Precalculus	✔	✔	✔			
Advanced Placement Statistics						–✔
An Advanced Placement calculus course	✔				✔	
Highest science course taken in high school (reference category is Chemistry)						
Biology				✔		✔
An Advanced Placement or International Baccalaureate biology course						
Integrated physics and chemistry						
Chemistry (reference category)						
Physics	✔		✔			
Advanced Placement Physics					✔	
Number of math and science courses taken in high school						
Number of math courses					✔	
Number of Advanced Placement math courses						✔
Number of science courses			✔		✔	
Number of Advanced Placement science courses	✔					
Standardized assessment score						
Grade 10 math	✔					
Grade 11 math	✔	✔	✔			
Grade 10 science					✔	
Grade 11 science	✔†	✔	✔			✔
SAT composite	—	—	—			

✔ is a significant positive association. –✔ is a significant negative association. † is less positive effects for Hispanic students (relative to non-Hispanic White students). — is not available.

STEM is science, technology, engineering, and math.

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) or a four-year Texas public or private college or university (a four-year college) by spring 2011. Courses are listed in the order of progression indicated by graduation plans specified by the Texas Education Agency.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

was students whose highest science course taken was Chemistry. (See appendix B for more information about checks for robustness, and see tables C1–C3 in appendix C for descriptive statistics for the full set of covariates in the models and tables C4–C7 in appendix C for descriptive statistics for the indicators disaggregated by race/ethnicity.)

Students' academic experiences in math and science in high school predicted declaring a STEM major. Among students who enrolled in a two-year college, academic experiences in math and science in high school predicted declaring a STEM major. Taking an advanced math course was strongly associated with the likelihood of declaring a STEM major. The most substantively significant finding was that students who took an Advanced Placement calculus course were 9.8 percentage points more likely to declare a STEM major than were students in the reference group (see table D1 in appendix D). Taking Precalculus was associated with a 2.7 percentage point increase in the likelihood of declaring a STEM major (relative to students in the reference group). Conversely, students whose highest math course was Algebra I, the lowest level course, were 0.7 percentage point less likely to declare a STEM major.

Achievement in math in high school also was associated with the likelihood of declaring a STEM major among students who enrolled in a two-year college. A one standard deviation increase in grade 10 state assessment score in math was associated with a 0.6 percentage point increase in the likelihood of declaring a STEM major (relative to students in the reference group), and a 1 standard deviation increase in grade 11 state assessment score in math was associated with a 1.4 percentage point increase in likelihood (see table D1 in appendix D).

Among students who enrolled in a two-year college, those who took an Advanced Placement calculus course were 9.8 percentage points more likely to declare a STEM major than were students whose highest math course taken was Algebra II

Courses taken in science and achievement in science in high school also had significant and robust positive associations with declaring a STEM major among students who enrolled in a two-year college. Students whose highest science course taken was Physics were 1.1 percentage points more likely to declare a STEM major than students in the reference group were (see table D1 in appendix D). The number of Advanced Placement science courses taken had a positive association with the likelihood of declaring a STEM major (relative to students in the reference group): each additional Advanced Placement science course taken was associated with a 2.5 percentage point increase in likelihood. Grade 11 state assessment score in science also had a positive association with the likelihood of declaring a STEM major: an increase of 1 standard deviation in the score was associated with a 1.2 percentage point increase in likelihood.

Among students who enrolled in a two-year college, a 1 percentage point increase in high school attendance rate was associated with a 0.1 percentage point increase in the likelihood of declaring a STEM major (see table D1 in appendix D).

Among students who enrolled in a four-year college, math and science courses taken in high school predicted declaring a STEM major. Contrary to expectations, students whose highest math or science course was the lowest in the course progression indicated by Texas Education Agency graduation plans had a greater likelihood of declaring a STEM major. Specifically, students whose highest math course taken was Algebra I were 4.2 percentage points more likely to declare a STEM major than students in the reference group were, and students whose highest science course taken was Biology were 6.2 percentage points more likely (see table D2 in appendix D).

These findings are counterintuitive because they indicate that not progressing past the lowest level of math or science course in high school is positively associated with declaring a STEM major among students who enrolled in a four-year college. One likely explanation: students who enrolled in a four-year college whose highest math course taken was Algebra I or whose highest science course taken was Biology took other math and science courses that were not captured in the taxonomy used to code courses (such as specialized courses with math applications relevant to specific careers; see appendix B for description of coding procedures) and that were not taken by students who enrolled in a two-year college. This explanation is consistent with the differences in the number of math and science courses required for admission to two- and four-year colleges in Texas.

Students' academic experiences in math and science in high school and high school attendance rate predicted persisting in a STEM major. Among students who enrolled in a two-year college, math courses taken and achievement in math and science in high school were positively and robustly associated with persisting in a STEM major. Students whose highest math course taken was Precalculus were 4.3 percentage points more likely to persist in a STEM major than students in the reference group were (see table D1 in appendix D). Grade 11 state assessment score in math had a positive association with the likelihood of persisting in a STEM major (relative to students in the reference group): a one standard deviation increase in score was associated with a 2.8 percentage point increase in likelihood (the largest effect size of all state assessment scores). A one standard deviation increase in grade 11 state assessment score in science was associated with a 1.4 percentage point increase in the likelihood of persisting in a STEM major.

Among students who enrolled in a four-year college, those whose highest math course taken was an Advanced Placement calculus course were 9.2 percentage points more likely to persist in a STEM major than were students whose highest math course taken was Algebra II

Among students who enrolled in a four-year college, math and science courses taken and achievement in math and science in high school predicted persisting in a STEM major. Students whose highest math course taken was an Advanced Placement calculus course were 9.2 percentage points more likely to persist in a STEM major than students in the reference group were, and students whose highest science course taken was Advanced Placement or International Baccalaureate Physics were 4.7 percentage points more likely (see table D2 in appendix D). The number of math courses taken and the number of science courses taken both had a positive association with the likelihood of persisting in a STEM major (relative to students in the reference group): each additional math course taken was associated with a 1.1 percentage point increase in likelihood, and each additional science course taken was associated with a 1.8 percentage point increase. Grade 10 state assessment score in science also had a positive association with the likelihood of persisting in a STEM major: a one standard deviation increase in score was associated with a 1.1 percentage point increase in likelihood.

Among students who enrolled in a four-year college, a 1 percentage point increase in high school attendance rate was associated with a 1.2 percentage point increase in the likelihood of persisting in a STEM major (see table D2 in appendix D).

Students' academic experiences in math and science in high school and high school attendance rate predicted completing a STEM degree. Among students who enrolled in a two-year college, math and science courses taken and achievement in math and science in high school predicted completing a STEM degree. Students whose highest math course taken was Precalculus were 1.4 percentage points more likely to complete a STEM degree than students in the reference group were, while students whose highest math course taken was Algebra I or Geometry were 0.3 percentage point less likely (see table D1 in appendix D). Students whose highest science course taken was Physics were 0.2 percentage point more likely to complete a STEM degree. The number of science courses taken had a positive association with the likelihood of completing a STEM degree: each additional science course taken was associated with a 0.2 percentage point increase in likelihood.

Among students who enrolled in a two-year college, grade 11 state assessment scores in math and science were positively associated with the likelihood of completing a STEM degree. A one standard deviation increase in math score was associated with a 0.7 percentage point increase in likelihood (relative to students in the reference group), and a one standard deviation increase in science score was associated with a 0.4 percentage point increase in likelihood (see table D1 in appendix D).

Among students who enrolled in a two-year college, a 1 percentage point increase in the high school attendance rate was associated with a 0.1 percentage point increase in the likelihood of completing a STEM degree (see table D1 in appendix D).

Among students who enrolled in a four-year college, math and science courses taken and achievement in science in high school predicted completing a STEM degree. The number of Advanced Placement math courses a student took was positively associated with the likelihood of completing a STEM degree (relative to students in the reference group): each additional course taken was associated with a 4 percentage point increase in likelihood (see table D2 in appendix D). However, students whose highest math course taken was Advanced Placement Statistics were 6.0 percentage points less likely to complete a STEM degree, perhaps because this small group (about 5 percent of the sample) took the course to prepare for a non-STEM major to which statistics would be relevant, such as business or a program in the social sciences. Students whose highest science course taken was the lowest level course in the taxonomy, Biology, were 3.7 percentage points more likely to complete a STEM degree than students in the reference group were. This result is unexpected because Biology is the first course in the science course progression; the reasons are likely similar to the reasons behind the similar finding for declaring a STEM major.

Among students who enrolled in a two-year college, math and science courses taken and achievement in math and science in high school predicted completing a STEM degree. Among students who enrolled in a four-year college, math and science courses taken and achievement in science in high school predicted completing a STEM degree

Grade 11 state assessment score in science was positively associated with the likelihood of completing a degree in a STEM field: a one standard deviation in score was associated with a 1.9 percentage point increase in likelihood (see table D2 in appendix D).

Several indicators did not predict postsecondary STEM success. Among students who enrolled in a two-year college, the total number of math courses taken and the number of Advanced Placement math courses taken in high school were not significantly and robustly associated with postsecondary STEM success.

Among students who enrolled in in a four-year college, grade 10 and 11 state assessment scores in math, SAT composite score, and the number of Advanced Placement science courses taken were not significantly and robustly associated with postsecondary STEM success.

Most indicators that predict postsecondary STEM success did not function differently for Hispanic students and for non-Hispanic White students

For the most part the indicators that predict postsecondary STEM success functioned similarly for Hispanic students and non-Hispanic White students.[3] The one exception was the relationship between grade 11 state assessment score in science and declaring a STEM major among students who enrolled in a two-year college. A one standard deviation increase in score was associated with a 1.5 percentage point increase in the likelihood of declaring a STEM major for non-Hispanic White students and a 1 percentage point increase for Hispanic students (see table D5 in appendix D).[4]

Implications of the study findings

This study found that indicators of student academic experiences, achievement in math and science, and high school attendance rate were associated with postsecondary STEM success. The total number and level of math and science courses taken and state assessment scores in math and science were similarly predictive of postsecondary STEM outcomes for Hispanic students and non-Hispanic White students. Of 29 statistically significant relationships between indicators and postsecondary outcomes, only 1 was qualified by an interaction with Hispanic/non-Hispanic White status. In other words, math and science courses taken and state assessment scores in high school are not differently associated with postsecondary STEM success for Hispanic students and for non-Hispanic White students.

Math and science courses taken and state assessment scores in high school are not differently associated with postsecondary STEM success for Hispanic students and non-Hispanic White students

A limited number of past studies have examined whether indicators performed differently for Hispanic students and for non-Hispanic students, and studies have not attempted to systematically examine the differences across a wide range of postsecondary STEM outcomes. The current study's systematic examination of the performance of indicators across postsecondary STEM outcomes yielded several findings that are consistent with those of past studies:

- *Number of STEM courses taken.* Three recent studies found that students who took more math and science courses in high school were more likely to declare a STEM major (Wang, 2013a, 2013b) and complete a STEM degree (Burge, 2013). The current study had similar findings; it also found a significant relationship between total number of math and science courses taken in high school and persisting in a STEM major among students who enrolled in a two-year college and among students who enrolled in a four-year college. Whereas Wang (2013b) found that number of math and science courses taken in high school performed strongest for non-Hispanic White students and weakest for underrepresented racial/ethnic minority students (Hispanic, non-Hispanic Black, and Native American students grouped together), the current study did not observe differences between Hispanic students and non-Hispanic students.
- *Number of Advanced Placement STEM courses taken.* Previous studies found that completing Advanced Placement STEM courses had a positive association with declaring a STEM major (Griffith, 2010; Shaw & Barbuti, 2010; You, 2013) and completing a STEM degree (Kokkelenberg & Sinha, 2010; Tyson, Lee, Borman, & Hanson, 2007). The current study also found a significant positive association between number of Advanced Placement math courses taken in high school and completing a STEM degree (among students who enrolled in a four-year college) and between number of Advanced Placement science courses taken and declaring a STEM major (among students who enrolled in two-year college).
- *Highest math course taken.* Previous studies found a strong relationship, which did not differ by race/ethnicity, between the highest math and science course taken in high school and postsecondary STEM success (Riegle-Crumb & King, 2010; You, 2013). The current study replicated those findings.
- *Student achievement.* Numerous past studies identified an association between achievement in math and science in high school and postsecondary STEM success (see Hinojosa et al., 2016). In the current study, among students who enrolled in

a two-year college, grade 11 state assessment scores in both math and science predicted declaring and persisting in a STEM major and completing a STEM degree. Science score was slightly (but significantly) less predictive of declaring a STEM major for Hispanic students than for non-Hispanic White students. There are several possible explanations for this finding—for example, that Hispanic students selected STEM majors for which science achievement in high school is not a strong prerequisite or that Hispanic students emphasized science achievement less than non-Hispanic White students did when selecting a two-year STEM major. Among students who enrolled in a four-year college, grade 10 state assessment score in science was positively associated with persisting in a STEM major, and grade 11 state assessment score in science was associated with completing a STEM degree.

Although seven past studies observed that students with higher SAT or ACT math scores were more likely to achieve postsecondary STEM success (see Hinojosa et al., 2016), the current study found no such association for composite SAT score (the sum of math and verbal subtest scores). The math subtest score was not available, and it seems likely that including the verbal subtest score in the composite score nullified the association of the math subtest score with the outcomes.

Increased participation in rigorous math and science courses in high school could reduce disparities in postsecondary STEM success between Hispanic students and non-Hispanic White students

In summary, this study demonstrates that Hispanic students reap the same benefits of taking higher level math and science courses in high school as non-Hispanic White students do in terms of postsecondary STEM outcomes. Among students who enrolled in a two-year college and students who enrolled in a four-year college, Hispanic students took fewer math and science courses overall, and took less rigorous courses, than non-Hispanic White students did (see table C4 in appendix C), which may explain the negative relationship between Hispanic status and postsecondary STEM outcomes. However, when indicators of academic experiences and achievement in math and science in high school were held constant, the likelihood of postsecondary STEM success did not differ between Hispanic students and non-Hispanic White students.

What contributes to Hispanic students taking fewer, and less rigorous, STEM courses in high school than non-Hispanic White students do? One possibility is that lower social or cultural capital available to Hispanic high school students restrains the formation of STEM identity—the ability to identify with STEM professions (DeWitt et al., 2011; see also Aschbacher, Li, & Roth [2010] for additional discussion of cultural influences). Future studies could consider these and other possible factors influencing the academic experiences of Hispanic students in math and science in high school.

More urgent implementation of policies and practices that promote rigor of courses offered as well as access to those courses could be considered. Increased participation in rigorous math and science courses in high school could reduce disparities in postsecondary STEM success between Hispanic students and non-Hispanic White students.

No high school indicator had a similar relationship to postsecondary STEM outcomes for students who enrolled in a two-year college and for students who enrolled in a four-year college. In some cases the differences are intuitive. For example, taking Precalculus as the highest math course was positively associated with postsecondary STEM success among students who enrolled in a two-year college but not among students who attended a four-year college. This most likely reflects the fact that students who enrolled in a two-year college typically took Algebra II as their highest course, so that taking Precalculus indicates a higher level of math learning than is typical, while for students who enrolled in a four-year college, taking Precalculus is more typical.[5] Other differences are less straightforward, such as the association between number of math courses (total and Advanced Placement) taken in high school and postsecondary STEM success among students who enrolled in a four-year college but not among students who enrolled in a two-year college, and the association between state assessment scores in math and postsecondary STEM success among students who enrolled in a two-year college but not among students who enrolled in a four-year college. Regardless of individual variables, the implications for the full set of indicators are similar for both groups of students: taking more and more-challenging math and science courses in high school is associated with more-positive postsecondary STEM outcomes.

Other differences unrelated to high school coursework and achievement—such as a student's ties to the community and family, status as a first-generation graduate, and perception of the value of postsecondary STEM education—may influence postsecondary STEM outcomes. Further investigation into why some students enroll in, persist in, and complete a postsecondary STEM degree could examine such individual factors.

In interpreting the findings, it is important to consider that the relationship between each indicator and postsecondary STEM success was examined while holding all other indicators constant

In interpreting the findings from the current study, it is important to consider that the relationship between each indicator and postsecondary STEM success was examined while holding all other indicators constant. This approach shows the independent relationship of each indicator to each outcome. However, when two indicators are correlated with each other (such as number of Advanced Placement math courses taken and state assessment score in math) holding one of the indicators constant can reduce the strength of the relationship between the other indicator and the outcome. For example, total number of science courses taken and grade 11 assessment score in science may be highly correlated with each other and with completing a STEM degree, but when total number of science courses taken is held constant, grade 11 assessment score in science may not have had a statistically significant association with the outcome, even though the unconditional association of grade 11 assessment score in science and completing a STEM degree is statistically significant. Therefore, the overall pattern of the findings is most important—that both the quantity and rigor of high school math and science courses taken in high school are associated with postsecondary STEM success.

Limitations of the study

The relationships that this study found between both fixed and malleable high school academic factors and postsecondary STEM outcomes are correlational; as a result, causal inferences cannot be made. Other factors may explain the statistical relationship. At best, a statistically significant relationship may suggest that a study with a stronger research design (such as a randomized controlled trial) could reveal a causal relationship.

This study included only indicators for which data were available in Texas's state longitudinal data system. Data on several potential predictors of postsecondary STEM outcomes are not captured in the system, including student grade point average and such noncognitive measures as students' attitudes toward math and science and their self-perceived ability in these areas. The study team coded courses by level but did not have access to any measure of course instructional or assessment practices or to related factors such as teacher capacity. Although Advanced Placement coursetaking and completion are available through the

Texas data system, Advanced Placement exam scores are not, so the study was unable to consider them even though research has shown that they are important correlates of postsecondary STEM outcomes (Hinojosa et al., 2016). Other factors, such as parent social capital, that have not yet been identified as indicators that predict postsecondary STEM success, could be influences as well (Ryan & Ream, 2016).

The study suggests that access to rigorous math and science courses, such as Advanced Placement calculus courses, is critical, but it did not examine access (although it examined factors associated with access, such as the proportion of students eligible for the federal school lunch program in a school). Because of the policy implications of unequal access to rigorous courses, the relationship between access and postsecondary outcomes could be explored in future research.

The study suggests that access to rigorous math and science courses, such as Advanced Placement calculus courses, is critical, but it did not examine access

This study team coded math and science courses taken in high school according to a taxonomy of seven courses in each subject (see appendix B). Students may have taken courses with extensive math or science content that did not fit into those taxonomies, and if so, variables associated with the highest course taken and total number of courses taken would not include all math- or science-related courses.

The study included only students who enrolled in a Texas college by spring 2011. This rule excluded students who did not enroll in college soon after high school.[6] The results for students in this excluded group may be systematically different from those for students who enrolled in college soon after completing high school. And the models in this study did not examine the characteristics of students who did not attend college shortly after high school or students who did not attend college at all; nor did the models include variables that might predict their decision to postpone or end their education.

Finally, the study did not examine differences within the Hispanic population (for example, between Hispanic students who are and those who are not eligible for the federal school lunch program or between those who are and those who are not proficient in English). The indicators may perform differently for these subgroups, so that combining them masks important distinctions. The different performance of the indicators among students who enrolled in a two-year college and students who enrolled in a four-year college underscores this possibility, a topic for future research.

Appendix A. Supplementary findings: Differences by sex and before and after a change in graduation requirements

This appendix reports supplementary findings on whether indicators that predict postsecondary STEM success function differently for female than for male students and before and after a change in state graduation requirements.

Do indicators function differently for female than for male students?

Among students who enrolled in a two-year college and students who enrolled in a four-year college, a lower percentage of female students than of male students achieved postsecondary science, technology, engineering, and math (STEM) success (table A1).

Female students were less likely than male students to declare and persist in a STEM major and complete a STEM degree. Even after differences in academic preparation were accounted for, female students were 11 percentage points less likely than male students to declare a STEM major among students who enrolled in a two-year college and 12 percentage points less likely among students who enrolled in a four-year college (see tables D1 and D2 in appendix D). The pattern was similar for persisting in a STEM major and completing a STEM degree. Female students were about 4 percentage points less likely than male students to persist in a STEM major among students who enrolled in a two-year college and 5 percentage points less likely among students who enrolled in a four-year college. Female students were about 1 percentage point less likely than male students to complete a STEM degree among students who enrolled in a two-year college and 2 percentage points less likely among students who enrolled in a four-year college. In other words, even when female students' STEM-related academic preparation is similar to that of male students and when female students' state assessment scores in math and science are similar to those of male students, female students are less likely than male students to declare a STEM major, persist in a STEM major, or complete a STEM degree.

Indicators of academic experiences in high school predict postsecondary STEM outcomes differently for female students and male students. This section highlights the significant interactions of indicators with a student's sex. Overall, the findings indicated a complex set of relationships among indicators, a student's sex, and postsecondary STEM outcomes.[7]

Table A1. Postsecondary STEM outcomes among Texas high school graduates, by type of college enrolled in and sex (percent)

	Students who enrolled in a two year college		Students who enrolled in a four year college	
Outcome	Male	Female	Male	Female
Declaring a STEM major	17.9	6.5	37.7	21.7
Persisting in a STEM major	44.5	40.7	59.0	50.2
Completing a STEM degree	2.5	1.1	11.4	6.5

STEM is science, technology, engineering, and math.

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) or a four-year Texas public or private college or university (a four-year college) by spring 2011. See appendix B for outcome definitions and a description of the analytic methodology.

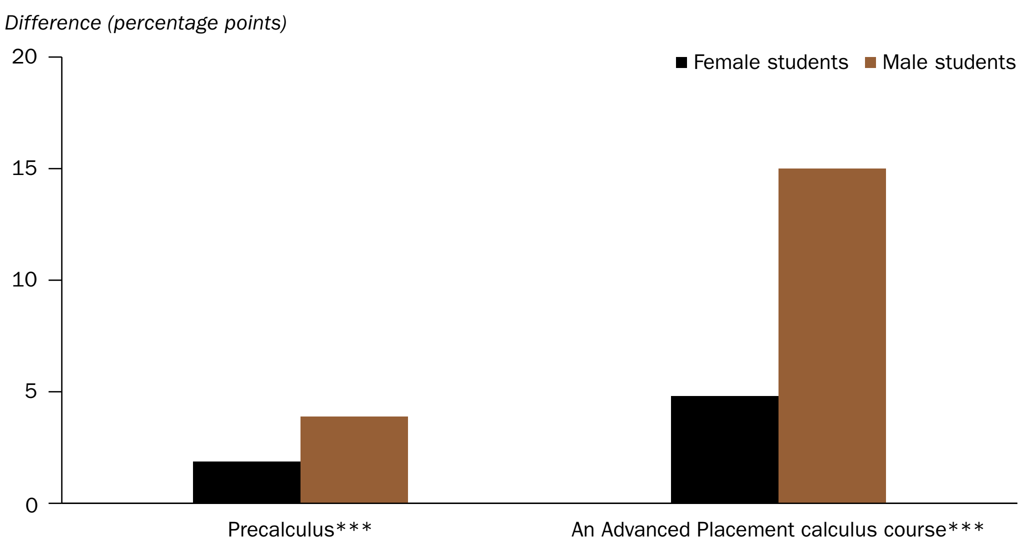
Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

The association between highest math course taken in high school and declaration of a STEM major differed between male and female students who enrolled in a two-year college. The interaction was most pronounced among students whose highest math course taken was an Advanced Placement calculus course (figure A1). Among students who enrolled in a two-year college, taking an Advanced Placement calculus course was associated with a 5 percentage point increase in the likelihood of declaring a STEM major (relative to someone whose highest math course was Algebra II) for female students and a 15 percentage point increase for male students. A similar but less dramatic difference was observed among students whose highest math course taken was Precalculus; among male students this highest math course was associated with a 4 percentage point increase in the likelihood of declaring a STEM major, and among female students it was associated with a 2 percentage point increase.

Among students who enrolled in a two-year college, grade 10 state assessment score in math was also less positively associated with declaring a STEM major for female students than for male students. A one standard deviation increase in score was associated with a third of a percentage point increase in the likelihood of declaring a STEM major for female students and a nearly 1 percentage point increase for male students.

Finally, the association between highest math course taken in high school and completing a STEM degree differed between male and female students who enrolled in a two-year

Figure A1. Among Texas high school graduates who enrolled in a two-year college, the association between an Advanced Placement calculus course as the highest math course taken and declaring a STEM major and between Precalculus as the highest math course taken and declaring a STEM major were both weaker for female students than for male students



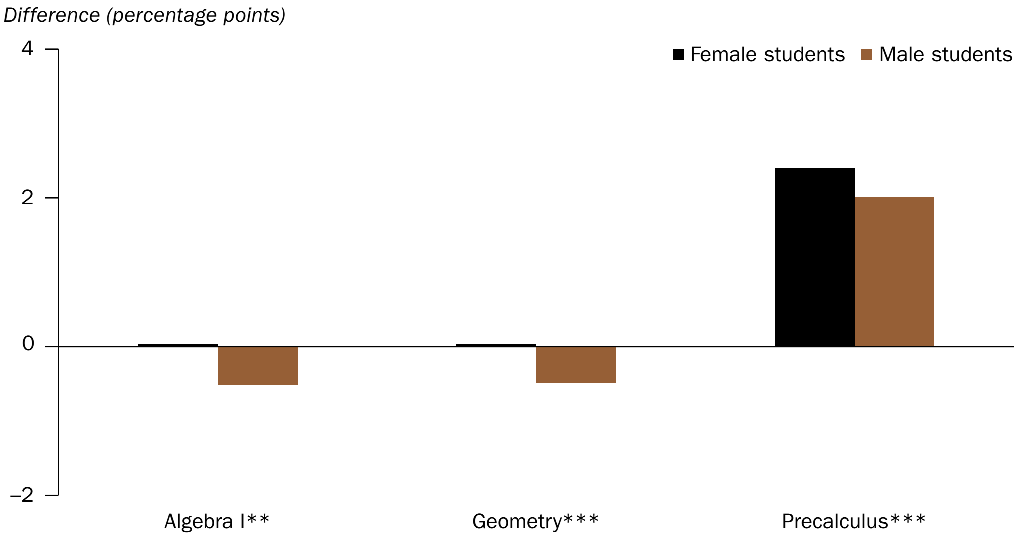
*** Statistically significant at $p < .001$.

STEM is science, technology, engineering, and math.

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) by spring 2011. Significance tests refer to the interaction term.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

Figure A2. Among Texas high school graduates who enrolled in a two-year college, having Algebra I, Geometry, or Precalculus as the highest math course taken had a less negative or more positive association with completing a STEM degree for female students than for male students



** Statistically significant at $p < .01$; *** statistically significant at $p < .001$.

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) by spring 2011. Significance tests refer to the interaction term.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

college, although it was more positive for female students than for male students. Having a relatively low-level math course (Algebra I or Geometry) as the highest math course taken was associated with a 0.5 percentage point decrease in the likelihood of completing a STEM degree for male students but a negligible increase for female students (figure A2). By the same token, among students who enrolled in a two-year college, taking Precalculus (a relatively high-level math course for this sample) was associated with a 2.4 percentage point increase in the likelihood of completing a STEM degree for female students and a 2.0 percentage point increase for male students.

Do indicators function differently before and after a change in state graduation requirements?

Beginning with the cohort that entered grade 9 in 2004/05, students had to complete Texas's Recommended High School Program to earn a diploma, unless the student, the student's parent or guardian, and a school counselor or school administrator agreed that the student should be permitted to take courses under the Minimum High School Program. To examine whether this policy influenced the relationship between indicators and postsecondary STEM outcomes, a modified version of model 2 included a dummy variable for whether a student entered grade 9 in 2004/05 or later (and was thus subject to the revised policy), as well as interactions between that dummy variable and the indicators (see appendix B).

Among students who enrolled in a two-year college, students in the three cohorts that were subject to the revised graduation policy were 2.2 percentage points less likely than students in cohorts that were not subject to the policy to declare a STEM major. Among students who enrolled in a four-year college, the difference was not statistically significant. The marginal effect of the policy on persisting in a STEM major is confounded by the smaller number of semesters of data in the cohorts subject to the revised policy and therefore is not interpretable.[8] Similarly, it was not possible to examine the association between the revised graduation policy and completing a STEM degree because there were too few semesters of data.

Indicators predicted declaring a STEM major differently for cohorts subject to the revised graduation policy and for cohorts not subject to the revised policy. Among students who enrolled in a two-year college, some indicators had a weaker association with declaring a STEM major for cohorts subject to the revised graduation policy than for cohorts not subject to it:

- Among cohorts not subject to the revised policy, students whose highest math course taken in high school was Precalculus were nearly 4 percentage points more likely to declare a STEM major than were students whose highest math course taken was Algebra II. Among cohorts subject to the revised policy, the likelihood increased 2.5 percentage points.
- Among cohorts not subject to the revised policy, students whose highest science course taken was Physics were about 1.7 percentage points more likely to declare a STEM major than were students whose highest science course taken was Chemistry. Among cohorts subject to the revised policy, the likelihood increased about 0.8 percentage point.
- Among cohorts not subject to the revised policy, a one standard deviation increase in grade 11 state assessment score in science was associated with a 1 percentage point increase in the likelihood of declaring a STEM major. Among cohorts subject to the revised policy, the increase in likelihood was 0.5 percentage point.

No statistically significant interactions of indicators with graduation policy were observed for declaring a STEM major for the four-year college sample.

Indicators predicted persisting in a STEM major differently for cohorts subject to the revised graduation policy and for cohorts not subject to the revised policy. Among students who enrolled in a two-year college, some indicators had a weaker association with persisting in a STEM major for cohorts subject to the revised graduation policy than for cohorts not subject to it:

- Among cohorts not subject to the revised policy, students whose highest math course taken in high school was Precalculus were about 6.5 percentage points more likely to persist in a STEM major than were students whose highest math course taken was Algebra II. Among cohorts subject to the revised policy, the likelihood increased 2.3 percentage points.
- Among cohorts not subject to the revised policy, a one standard deviation increase in grade 11 state assessment score in math was associated with a nearly 3 percentage point increase in the likelihood of persisting in a STEM major. Among cohorts subject to the revised policy, the likelihood increased 0.5 percentage point.
- Among cohorts not subject to the revised policy, a one standard deviation increase in grade 11 state assessment score in science was associated with a 1.7 percentage point increase in the likelihood of persisting in a STEM major. Among cohorts subject to the revised policy, the likelihood increased about 0.3 percentage point.

Among students who enrolled in a four-year college, the revised graduation policy influenced the association between having an Advanced Placement calculus course as the highest course taken and persisting in a STEM major. Among cohorts not subject to the revised policy, students who took an Advanced Placement calculus course were about 11 percentage points more likely to persist in a STEM major than were students whose highest math course taken was Precalculus. Among cohorts subject to the revised policy, the likelihood increased about 5 percentage points.

Appendix B. Data sources, sample, and methodology

This appendix describes the data sources from which the study dataset was constructed, the sample of students created from these sources, the variables created from the available data, and the methods for analyzing those variables.

Data sources

All the data analyzed in this study were accessed through Texas's state longitudinal data system, which contains a broad range of student- and school-level data from a variety of sources, including the Texas Education Agency and the Texas Higher Education Coordinating Board, that can be linked with students and over time. Housed in Education Research Centers across the state, the system includes data on students' demographic characteristics, K–12 achievement and assessment scores, courses taken, and postsecondary outcomes (that is, college enrollment, major, and degree completion). The study team used the following records:

- *Student-level accountability records reported to the Texas Education Agency by school districts.* Texas school districts are required to report data on each student to the Texas Education Agency at several points throughout the year. These data are stored in the agency's Public Education Information Management System database and then incorporated into the Education Research Center database. Only public schools, including open-enrollment charter schools, are included. The study used these Public Education Information Management System records:
  - Student demographic characteristics (race/ethnicity, eligibility for the federal school lunch program, sex, and home language).
  - Student classification for special status (special education classification and English learner status).
  - Student coursetaking records (enrollment in math and science courses).
  - Student attendance.
  - Student graduation from Texas public schools.
- *Student-level state assessment records collected by the Texas Education Agency.* The study team accessed students' math and science state assessment records from the state's standardized assessment data in the Education Research Center database. Two assessment regimes were used during the period covered by the study: the Texas Assessment of Academic Skills, which was phased out in the 2002/03 school year, though some students may have been subject to the exit-level test requirements after 2003, and the Texas Assessment of Knowledge and Skills, which was deployed in 2003. The study team used valid scores only from the first administration of each test. The analyses used standardized scores based on the pooled, within-year, grade, and subject-area test standard deviation.
- *Postsecondary enrollment records collected by the Texas Higher Education Coordinating Board.* The study team linked data from students' high school academic experiences to college enrollment data provided by the Texas Higher Education Coordinating Board on students who enrolled in a college or university in Texas. These postsecondary records were available from three types of institutions of higher education: two-year public, four-year public, and four-year private. Two-year Texas public colleges are referred to as two-year colleges in the analyses, and four-year Texas public and private colleges and universities are referred to as four-year colleges in the analyses. The records provided to the Education Research Center

by the Texas Higher Education Coordinating Board included student-level postsecondary enrollment records for all Texas students attending public or private institutions of higher education. These records contained the following student data used in this study to construct the outcomes:

- Enrollment history.
- Declared major (available only for two- and four-year public institutions of higher education, not four-year private institutions of higher education).
- Graduation degree.
- Degree major.
- SAT composite math and verbal subtest scores.

Information on student majors was available only for public institutions, so only students who enrolled in a public institution have data for the outcomes of declaring and persisting in a STEM major.

For each covariate for which data were missing, the study team assessed whether there was a substantively significant difference in the mean values of the covariate between the analytic sample (cases with no missing data for covariates) and the full sample (cases with missing data for at least one covariate). The criteria to determine whether a difference was substantively significant followed Little and Rubin's guidelines (2002). In this approach the difference in a covariate's mean value between the full and the analytic sample is considered substantively significant when it is greater than 0.1 standard deviation or less than –0.1 standard deviation. When the standardized difference in the mean for a covariate was substantively significant, a dummy variable imputation method was used to prevent a loss of observations that would have created covariate imbalance between the full and analytic samples. Dummy variable imputation is especially appropriate for analyzing large datasets, such as the one used in this study, for which it is likely that most or all differences will be statistically significant. If the standardized difference was not substantively significant, case deletion was used instead.

Missing observations that met the substantive significance criterion for a covariate were imputed with the mean of the covariate. Then, a dummy variable for each covariate was created with a value of 1 for observations that were missing and 0 otherwise. Covariates for which missingness was associated with substantive differences between the full and analytic sample were highest math course taken in high school, highest science course taken in high school, and grade 10 and 11 state assessment scores in math and science for students who enrolled in a two-year college and SAT composite score, highest math course taken in high school, and highest science course taken in high school among students who enrolled in a four-year college.

Cohort membership

Cohort membership was determined by finding the first year a student was in grade 9 between 2000/01 (cohort 1) and 2006/07 (cohort 7; table B1). Cohorts were not fixed, and students were not required to have a grade 9 record. Rather, students who arrived after grade 9 were folded into the cohort corresponding to their grade–school year configuration and to the point in each cohort's grade progression between grade 9 and grade 12. For example, cohort 1 includes students who were in grade 9 in the 2000/01 school year. However, a student who first appeared in Texas Education Agency's enrollment records

Table B1. Cohorts of Texas high school graduates included in the study and STEM outcome data available by year, 2000/01–2010/11

Cohort	2000/01	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09	2009/10	2010/11
1	Grade 9	Grade 10	Grade 11	Grade 12	Declare	Declare Persist Degree[a]	Declare Persist Degree[a]	Declare Persist Degree	Declare Persist Degree	Declare Persist Degree	Declare Persist Degree
2	—	Grade 9	Grade 10	Grade 11	Grade 12	Declare	Declare Persist Degree[a]	Declare Persist Degree[a]	Declare Persist Degree	Declare Persist Degree	Declare Persist Degree
3	—	—	Grade 9	Grade 10	Grade 11	Grade 12	Declare	Declare Persist Degree[a]	Declare Persist Degree[a]	Declare Persist Degree	Declare Persist Degree
4	—	—	—	Grade 9	Grade 10	Grade 11	Grade 12	Declare	Declare Persist Degree[a]	Declare Persist Degree[a]	Declare Persist Degree
5	—	—	—	—	Grade 9	Grade 10	Grade 11	Grade 12	Declare	Declare Persist Degree[a]	Declare Persist Degree[a]
6	—	—	—	—	—	Grade 9	Grade 10	Grade 11	Grade 12	Declare	Declare Persist Degree[a]
7	—	—	—	—	—	—	Grade 9	Grade 10	Grade 11	Grade 12	Declare

STEM is science, technology, engineering, and math. Declare is declare a STEM major. Persist is persist in a STEM major. Degree is complete a STEM degree.

a. Degree data are for associate's degrees only.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

in grade 10 in 2001/02 was added to cohort 1. Students who did not progress sequentially through high school (for example because they repeated or skipped a grade) were not reassigned to a different cohort. Cohort assignment was permanent.

Definition of variables

This section provides additional detail on the definition of the outcomes and indicators. The three measures of postsecondary STEM outcomes were binary (yes/no) variables:

- Declaring a STEM major: whether a college student ever declared a STEM major.
- Persisting in a STEM major: whether a college student who declared a STEM major remained enrolled as a STEM major for all subsequent semesters. This variable was calculated for students who had more than one semester of data, who declared a STEM major at some point, and who had at least one subsequent semester of data. Students with breaks in their postsecondary experience were not excluded, as long as they returned to college and continued their STEM major.
- Completing a STEM degree: whether a college student completed a degree (even from a different institution of higher education than at first enrollment) with a STEM major. The analyses of students who enrolled in a two-year college examined whether they completed a two-year or a four-year STEM degree, and the analyses of students who enrolled in a four-year college examined whether they completed a four-year STEM degree (regardless of whether they previously or subsequently completed a two-year STEM degree). No time limit on graduation was applied; students could have been coded as having completed a STEM degree in any school year after enrollment.

Each outcome measure was analyzed separately for students who enrolled in a two-year college and students who enrolled in a four-year college. Although students who transferred from a two-year college to a four-year college (or vice-versa) were retained, assignment to the two-year college sample or the four-year college sample was permanent. Thus, students who enrolled in a two-year college and transferred to and declared a major at a four-year college were grouped with other students who enrolled in and declared a major at a two-year college.[9]

Constructing the outcomes required categorizing the major as belonging to a STEM field or not. Following the example of the National Center for Education Statistics Web Tables, STEM fields were defined as computer and information sciences, engineering and engineering technologies, biological and biomedical sciences, math and statistics, physical sciences, and science technologies (National Center for Education Statistics, 2011).

Several indicators described experiences with math and science courses in high school. High school courses were categorized as math or science on the basis of courses listed in graduation credit requirements in Texas (Texas Education Agency, 2012).[10] Three types of variables were created from course records, separately for math and science: the total number of courses taken, the number of Advanced Placement courses taken, and the highest course taken. Students with only three years of high school records in the dataset provided by Texas Education Agency were expected to have lower values for total number of courses taken, so a variable was included in the statistical model for number of years enrolled in a Texas public high school.

The indicator highest course taken reflected the order of progression indicated in graduation plans specified by the Texas Education Agency. Math courses were ordered:

- Algebra I.
- Math Models.
- Geometry.
- Algebra II.
- Precalculus.
- Advanced Placement Statistics.[11]
- An Advanced Placement calculus course.[12]

Science courses were ordered:

- Biology.
- Advanced Placement Biology.
- Integrated Physics and Chemistry.
- Chemistry.
- Advanced Placement Chemistry.
- Physics.
- Advanced Placement Physics.

Courses that did not fit into these taxonomies were not coded as a math or science course. For example, a highly specialized course with math applications relevant to a specific career would not be counted toward the total number of math courses.

All variables used in the study are summarized in table B2.

Table B2. Variables used in the study: Outcomes, indicators, school-level contextual variables, student-level covariates, and control variables

Variable	Operationalization
Outcome[a]	
Declaring a STEM major	Whether a student declared a major that was categorized as a STEM major.
Persisting in a STEM major	Whether a student who declared a STEM major remained a STEM major for all subsequent semesters.
Completing a STEM degree	Whether a student earned a degree that was categorized as a STEM major.
Indicator[b]	
Indicators identified in the literature review	
SAT composite score	Sum of SAT math and verbal subtest scores. Subtest scores were not available in the Public Education Information Management System.
Total number of math courses taken	Total number of math courses taken in high school.
Total number of science courses taken	Total number of science courses taken in high school.
Number of Advanced Placement math courses taken[c]	Total number of Advanced Placement math courses taken in high school.
Number of Advanced Placement science courses taken[c]	Total number of Advanced Placement science courses taken in high school.
Highest math course taken	Dichotomous variables for each math course. Students were categorized by the most advanced course taken according to the Texas graduation plan course progression. Algebra II was the largest category and therefore served as the reference group.[d]
Highest science course taken	Dichotomous variables for each science course. Students were categorized by the most advanced course taken according to the Texas graduation plan course progression. Chemistry was the largest category and therefore served as the reference group.
State assessment scores in math and science	Math and science scores on state assessments (Texas Assessment of Academic Skills or Texas Assessment of Knowledge and Skills, depending on the year) administered in grades 10 and 11.
Exploratory indicator	
High school attendance rate	The percentage of days attended relative to days enrolled (in the regression model, this variable was centered at the overall mean).
School-level contextual variable[e,f]	
Percentage of students in an English as a second language program	Percentage of students in a student's high school who were in an English as a second language program.
Racial/ethnic composition of the school	Three variables were included for the proportion of students in a student's school who are Hispanic, non-Hispanic Black, and other (American Indian or Alaska Native, Asian, or Native Hawaiian or Other Pacific Islander). Non-Hispanic White is the reference group.
Percentage of students eligible for the federal school lunch program	Percentage of students in a student's high school eligible for the federal school lunch program.
Student-level covariate[e]	
Race/ethnicity	Dichotomous variables for whether a student is Hispanic, non-Hispanic Black, or other (American Indian or Alaska Native, Asian, or Native Hawaiian or Other Pacific Islander). Non-Hispanic White is the reference group.
Sex	Dichotomous variable for whether a student is female.
Eligibility for the federal school lunch program	Dichotomous variable for whether the student was classified as eligible for the federal school lunch program at any time.
Special education	Dichotomous variable for whether the student was in special education at any time.
English learner status	Dichotomous variable for whether the student was classified as an English learner at any time.
Home language	Dichotomous variables for whether the student's primary home language was Spanish or other (neither English nor Spanish) at any time. English is the reference group.
Cohort	Dummy variables indicating the cohort of which the student is a member.

(continued)

Table B2. Variables used in the study: Outcomes, indicators, school-level contextual variables, student-level covariates, and control variables *(continued)*

Variable	Operationalization
Control variable	
Whether high school diploma was earned	Dichotomous variable for whether the student earned a high school diploma in Texas.
Number of years enrolled high school	Number of years in which a student was enrolled in a Texas public high school.

STEM is science, technology, engineering, and math.

a. All outcomes were obtained from Texas Higher Education Coordinating Board data, each file of which includes a variable for student's major.

b. All indicators were obtained from Public Education Information Management System with the exception of state assessment data, which were obtained from Texas Education Agency testing data.

c. Advanced Placement exam scores are not available in the Education Resource Center databases.

d. To allow the same models to be used for students who enrolled in a two-year college and students who enrolled in a four-year college, the same reference group was used for both groups, even though Precalculus was the largest group among students who enrolled in a four-year college.

e. Measured during the first high school year for which data on the student are available.

f. School-level contextual variables and student-level covariates were obtained from Public Education Information Management System.

Source: Authors' compilation.

Inclusion rules for the analytic sample

The initial sample included all students who:

- Enrolled in grade 9 or later in a Texas public school from 2000/01 (cohort 1) to 2005/06 (cohort 7).
- Were enrolled on the fall accountability snapshot date (for example, the last Friday in October) in each school year.
- Had a valid ID assigned through the Public Education Information Management System (those without a valid ID could not be reliably linked across school years in the Texas Education Agency records).

Three inclusion rules winnowed the initial sample to the analytic sample (table B3):

- *Rule 1:* at least one year of high school course enrollment data and a valid social security number. Matching across state agencies was performed using an anonymized social security number replacement identifier assigned by the Texas Higher Education Coordinating Board. Students without a valid social security number in the Texas Education Agency records could not be matched to postsecondary records. Students were required to have at least one year of high school enrollment data (in grades 9–12) because these data were the source for the indicators being studied.
- *Rule 2:* three or four years of valid public high school records. This rule minimizes the confounding effect of high school enrollment duration and the number and level of courses taken. It allows nonsequential grade trajectories for students: a student who enrolled in grade 9 in 2000/01, repeated grade 9 in 2001/02, was promoted to grade 10 in 2002/03, and dropped out without graduating in 2003/04 would remain in the cohort because she had three years of high school records.
- *Rule 3:* enrolled in a Texas institution of higher education by spring 2011. This rule had a differential impact on cohorts 6 and 7 because students who delayed postsecondary entry (for example, by taking a gap year) were not included in the student count.

Table B3. Impact of inclusion rules on the number of students in the sample, by cohort

Assigned cohort	Initial sample	Rule 1	Rule 2	Rule 3
1 (2000/01)	351,995	309,156	256,660	149,583
2 (2001/02)	346,542	304,311	258,482	148,854
3 (2002/03)	353,634	310,213	263,771	148,067
4 (2003/04)	361,360	316,975	269,071	147,528
5 (2004/05)	369,368	325,101	276,609	148,166
6 (2005/06)	375,885	330,768	281,358	143,842
7 (2006/07)	379,324	332,302	284,138	127,567

Note: Rule 1: at least one year of high school course enrollment data and a valid social security number. Rule 2: three or four years of valid public high school records. Rule 3: enrolled in a Texas institution of higher education by spring 2011.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

Comparing the samples from rule 2 to rule 3 indicates that, across the seven cohorts included in the study, 54 percent of Texas public high school graduates enrolled in a Texas institution of higher education by spring 2011. These students are the subject of this study. Table B4 summarizes the availability of data for each postsecondary STEM outcome as a function of cohort and type of college and summarizes the sample size for each outcome for students who enrolled in a two-year college and students who enrolled in a four-year college.

Analytic methodology

To identify indicators that predict declaring a STEM major, persisting in a STEM major, and completing a STEM degree, a linear probability model was fit for each of the binary

Table B4. Cohorts covered, data availability, and sample size, by type of college enrolled in and postsecondary STEM outcome among Texas high school graduates

		Type of college enrolled in			Sample size	
Outcome	Cohorts	Two year	Four year public	Four year private	Students who enrolled in a two year college	Students who enrolled in a four year college
Declare a STEM major	1–7	✔	✔	—	649,217	376,741
Persist in a STEM major	1–6	✔	✔	—	62,406[a]	102,944[a]
Earn a STEM degree (students who enrolled in a two-year college)	1–6	✔	na	na	649,217	na
Earn a STEM degree (students who enrolled in a four-year college)	1–4	—	✔	✔	—	455,560

— is not available. na is not applicable.

STEM is science, technology, engineering, and math.

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) or a four-year Texas public or private college or university (a four-year college) by spring 2011.

a. Includes only students who declared a STEM major and had one subsequent semester of data.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

outcomes of interest. The study team calculated robust standard errors to account for the heteroscedasticity that originates from applying ordinary least squares on a binary outcome (Wooldridge, 2002).[13] The calculation was performed in Stata using the robust option in each regression specification, which returns White-corrected standard errors (StataCorp, 2013, 20.21, p. 49). More precisely, this method inflates the standard errors for each coefficient estimate to account for violations of the assumption of independence between residuals *e* and a given covariate *x*—that is, heteroscedasticity—thereby reducing the likelihood of erroneously rejecting the null hypothesis, or committing a Type I error. The study team further adjusted standard errors to account for the nonindependence of observations—and, consequently, the correlation of the disturbances within high schools by clustering at the high school level—using the variance-covariance matrix adjustment cluster option with student's last high school of attendance (StataCorp, 2013, 20.21, p. 52).

The study team performed this modeling approach separately for each of the three outcomes. The coefficients derived from the linear probability model approximates the average marginal effect, expressed in probabilities, of a one standard deviation increase in x_k.

The full model includes a series of indicators modeled simultaneously and a series of school-level contextual variables and student-level covariates and takes the following form:

$$P(y = 1 \mid x) = \pi_0 + \sum_{p=1}^{P} \pi_a a_p + \sum_{q=1}^{Q} \lambda_q b_q + \sum_{k=1}^{K} \gamma_k c_k$$

where *y* is one of the three binary outcomes assessed for students in the analytic sample, a_p are the indicators,[14] c_k are school-level contextual variables, and b_q are student-level demographic covariates, as well as simple controls (student-level covariates that are assumed to have some association with the outcomes but are not demographic characteristics: whether a student earned a high school diploma and the number of years a student was enrolled in a Texas high school; see table B2).

Model comparisons. A systematic model comparison approach was employed to address the two research questions. This approach involved fitting four models that included indicators and student-level demographic covariates (table B5).

Table B5. Composition of regression models used to make model comparisons

	Models			
Variable	1	2	3	4
Indicators		✔	✔	✔
Student-level demographic covariates	✔	✔	✔	✔
School-level contextual variables	✔	✔	✔	✔
Student Hispanic status and the interactions between indicators and student Hispanic status				✔
Interactions between student Hispanic status and student-level demographic covariates, as well as school-level contextual variables			✔	✔

Source: Authors' compilation.

These models were used to conduct two main comparisons to identify indicators that are associated with outcomes (research question 1) and indicators that are associated with outcomes differently depending on whether a student is Hispanic (research question 2).

The analytic approach used Wald tests of whether blocks of additional covariates from the full model (such as indicators) improve the amount of variance explained over the reduced, nested models (Greene, 2003). The procedure involved fitting the nested[15] regressions sequentially and performing a series of significance tests to assess the constraint that the effects from each additional block of indicators are jointly different from zero.

Comparing model 1 with model 2. This comparison assessed whether including individual indicators increased the predictiveness of a model that included school-level contextual variables and student-level demographic covariates.

If the indicators were jointly significant (in terms of the Wald tests), the statistical significance of each indicator in the model was examined. For an indicator to be considered a demonstrated indicator, the set of indicators must be jointly significant (in terms of the Wald test), and the indicator must be statistically significant in the full model.

Given the premise that indicator estimates should be robust to inclusion of stable and time-invariant factors, comparing the two models allowed assessment of which indicators were significantly associated with outcomes even after the effects of the time-invariant covariates at the student level were controlled for.

Comparing model 3 with model 4. This comparison paralleled the first comparison by assessing whether interactions between Hispanic status and student-level covariates were jointly significant in a model that also included the indicators, school-level contextual variables, student-level demographic covariates, and interactions between Hispanic status and student-level covariates and school-level contextual variables.

After confirming that interactions between Hispanic status and student-level covariates were jointly significant (in terms of the Wald tests), it was possible to examine the statistical significance of the interaction between each indicator and Hispanic student status.

An indicator was considered to have different predictiveness for Hispanic students and non-Hispanic students under the following conditions: the interactions between Hispanic status and the indicator were jointly significant, the indicator was significant in model 3, and the interaction was significant in model 4.

The comparison of models 3 and 4 allowed assessment of factors that were differently predictive of outcomes for Hispanic students and non-Hispanic students even after differences in distributions between the two groups on stable factors (nonindicators) that were associated with the outcome or interacted with Hispanic status were accounted for.

Interpreting significant interaction terms. The study team sought to describe and interpret interactions that were both significant and robust to cohort and model composition. To this end, the study team added the beta for each interaction term to the beta for its constituent term and then multiplied the sum by 100. The resulting number is the differential effect of the indicator on Hispanic (versus non-Hispanic White) students, expressed as a percentage point difference. For example, the beta for the interaction term Hispanic * highest science course taken: Chemistry was 0.036. The beta of the constituent chemistry term was –0.045. The sum of these betas is –0.009, which, when multiplied by 100, indicates that Hispanic students whose highest science course taken was Chemistry were 0.9 percentage point less likely to complete a STEM degree than were Hispanic students whose highest science course taken was Biology. This negative effect is smaller than for non-Hispanic White students, who were 4.5 percentage points less likely to complete a four-year STEM degree if their highest science course taken was Chemistry.

Assessing the robustness of estimates of the final models. To ensure that conclusions are derived from associations between indicators and outcomes and not from other factors excluded from the models, the strength of the associations between indicators and outcomes was tested to be robust to:

- Changes of covariates used in the analytic models.
- Choice of study cohort or sample.

To address these two areas, three tests of robustness were applied to models 3 and 4 (discussed previously and summarized in table B5). The three robustness checks are discussed in more detail below (table B6). Only associations that were robust across these three tests were included in the findings.

Changes of covariates used in the analytic models. Indicator variables are malleable factors that are different from stable attributes of students or contextual factors. To declare a variable to be an indicator, it is important to demonstrate that the association between the indicator and any given outcome is not due to a confounding of the variable with student stable attributes or contextual factors. Because of the nonexperimental nature of an indicator analysis, full assurance of the nature of the relationship between the indicator and the outcome is not possible, but it is possible to control for the effects of other observable variables to a limited extent. To that end, for each indicator, high school fixed effects were added to the full model, and the statistical significance of the indicator effects in the presence of high school fixed effects was assessed. In this model, all independent variables were identified by variation within a high school as opposed to between schools (that is, students are compared with other students within their own high school).

Choice of study cohort. To be a reliable predictor of student outcomes, the effects of an indicator should be robust to different samples. Because the study sample is large, alternative sampling scenarios were used to test robustness. Two analyses were conducted

Table B6. Summary of tests of robustness applied to full and full Hispanic models

Robustness check	Model specification	Model details
1	High school fixed effects	Inclusion of high school fixed effects in the full and full Hispanic models
2	Student cohort effect	Interaction between student cohort and each indicator
3	Diploma requirement cohort effect	Interaction between 2004/05 or later high school entry indicator (demarcating the start of new diploma requirements) and each indicator

Source: Authors' compilation.

separately for each outcome, one examining a student cohort effect and one examining a changed diploma requirement student cohort effect (see appendix A).

Student cohort effect. This analysis included a cohort effect for students in order to understand whether the predictive power of the indicators for STEM success vary over time. There are seven grade cohorts, and four of them have data spanning grade 9 through college graduation.

Revised diploma requirement student cohort effect. This robustness test included a dummy variable identifying cohorts who entered grade 9 in 2004/05 or later, which demarcates the start of the state's revised diploma requirements. Beginning in that year, students had to complete Texas's Recommended High School Program to earn a diploma. Students could graduate under the Minimum High School Program requirements but had to receive permission from a parent or guardian and a school counselor or administrator. It is possible that the policy influenced patterns of graduating and subsequent declaring of a STEM major and changed the capacity of indicators to predict outcomes. The stability of indicator effects under the change of policy was examined by modeling the interaction between the dummy variable representing when the policy was present and the indicators.

Appendix C. Descriptive statistics for indicators

This appendix summarizes descriptive statistics (means, standard deviations, and percentages of missing data) for the variables included in the regression models.

Table C1. Descriptive statistics for covariates in the statistical models, Texas high school graduates who enrolled in a two-year college

Variable	Mean	Standard deviation	Percent missing
High school attendance rate[a] (percent)	–0.44	4.63	0.0
Grade 10 state assessment scaled score in math	–0.10	0.80	22.9
Grade 11 state assessment scaled score in math	–0.17	0.81	17.8
Grade 10 state assessment scaled score in science	–0.09	0.83	23.0
Grade 11 state assessment scaled score in science	–0.17	0.83	18.0
Number of Advanced Placement math courses taken	0.05	0.23	0.0
Number of Advanced Placement science courses taken	0.05	0.24	0.0
Total number of math courses taken	2.82	1.09	0.0
Total number of science courses taken	2.76	0.90	0.0
SAT composite score	917.8	170.4	87.6
Percentage of students who are English learner students[a]	0.38	6.6	0.0
Percentage of students who are eligible for the federal school lunch program[a]	3.19	24.9	0.0
Percentage of students who are non-Hispanic White[a]	–2.43	29.4	0.0
Percentage of students who are non-Hispanic Black[a]	0.45	16.0	0.0
Percentage of students who are non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander[a]	–0.40	4.6	0.0
Number of years enrolled in high school	3.9	0.4	0.0

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) by spring 2011. See appendix B for variable definitions.

a. Data are centered values.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C2. Descriptive statistics for covariates in the statistical models, Texas high school graduates who enrolled in a four-year college

Variable	Mean	Standard deviation	Percent missing
High school attendance rate[a] (percent)	1.59	2.74	0.0
Grade 10 state assessment scaled score in math	0.71	0.97	15.8
Grade 11 state assessment scaled score in math	0.63	0.96	5.0
Grade 10 state assessment scaled score in science	0.65	0.93	16.1
Grade 11 state assessment scaled score in science	0.57	0.95	5.6
Number of Advanced Placement math courses taken	0.36	0.59	0.0
Number of Advanced Placement science courses taken	0.29	0.60	0.0
Total number of math courses taken	3.49	0.90	0.0
Total number of science courses taken	3.31	0.84	0.0
SAT composite score	1,064.9	182.8	31.6
Percentage of students who are English learner students[a]	0.33	5.75	0.0
Percentage of students who are eligible for the federal school lunch program[a]	–2.01	24.3	0.0
Percentage of students who are non-Hispanic White[a]	0.81	28.9	0.0
Percentage of students who are non-Hispanic Black[a]	0.27	16.8	0.0
Percentage of students who are non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander[a]	0.60	5.8	0.0
Number of years enrolled in high school	4.0	0.2	0.0

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a four-year Texas public or private college or university (a four-year college) by spring 2011. See appendix B for variable definitions.

a. Data are centered values.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C3. Demographic characteristics of and highest math and science courses taken by Texas high school graduates, by type of college enrolled in

Covariate	Students who enrolled in a two year college		Students who enrolled in a four year college	
	Percent	Percent missing data	Percent	Percent missing data
Ever an English learner student	5.3	0.0	1.5	0.0
Ever eligible for the federal school lunch program	50.6	0.0	31.9	0.0
Ever in an English as a second language program	4.3	0.0	1.2	0.0
Ever in special education	11.3	0.0	2.4	0.0
Non-Hispanic White	45.7	0.0	53.8	0.0
Hispanic	38.2	0.0	25.7	0.0
Non-Hispanic Black	13.1	0.0	14.0	0.0
Other race/ethnicity[a]	2.9	0.0	6.5	0.0
Male	48.1	0.0	45.9	0.0
Female	51.9	0.0	54.1	0.0
Received a high school diploma	92.1	0.0	99.1	0.0
Highest math course taken				
Algebra I	3.3	3.9	0.5	0.4
Math Models	3.0	3.9	0.3	0.4
Geometry	16.0	3.9	3.4	0.4
Algebra II	53.1	3.9	28.9	0.4
Precalculus	20.0	3.9	37.2	0.4
Advanced Placement Statistics	1.2	3.9	5.2	0.4
An Advanced Placement calculus course	3.3	3.9	24.6	0.4
Highest science course taken				
An Advanced Placement or International Baccalaureate biology course	0.0	2.1	0.0	0.2
Biology	3.6	2.1	0.6	0.2
Integrated Physics and Chemistry	17.2	2.1	2.8	0.2
Chemistry	51.3	2.1	36.2	0.2
An Advanced Placement or International Baccalaureate chemistry course	0.4	2.1	1.7	0.2
Physics	26.5	2.1	50.9	0.2
An Advanced Placement or International Baccalaureate physics course	1.0	2.1	7.8	0.2

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) or a four-year Texas public or private college or university (a four-year college) by spring 2011. See appendix B for variable definitions.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C4. Descriptive statistics for indicators that predict postsecondary science, technology, engineering, and math success among Texas high school graduates who enrolled in a two-year college, by race/ethnicity

Variable		Non Hispanic Black	Hispanic	Non Hispanic White	Other[a]
High school attendance rate[b] (percent)	Mean	–0.45	–0.72	–0.27	0.65
	Standard deviation	4.89	4.85	4.36	4.32
Grade 10 state assessment scaled score in math	Mean	–0.51	–0.24	0.10	0.15
	Standard deviation	0.68	0.74	0.80	0.84
Grade 11 state assessment scaled score in math	Mean	–0.59	–0.30	0.02	0.07
	Standard deviation	0.69	0.75	0.81	0.84
Grade 10 state assessment scaled score in science	Mean	–0.52	–0.31	0.19	0.03
	Standard deviation	0.72	0.75	0.82	0.82
Grade 11 state assessment scaled score in science	Mean	–0.56	–0.37	0.09	–0.04
	Standard deviation	0.73	0.75	0.83	0.82
Number of Advanced Placement math courses taken	Mean	0.02	0.04	0.06	0.10
	Standard deviation	0.15	0.22	0.25	0.35
Number of Advanced Placement science courses taken	Mean	0.02	0.04	0.06	0.10
	Standard deviation	0.17	0.21	0.26	0.35
Total number of math courses taken	Mean	2.62	2.80	2.87	2.97
	Standard deviation	1.16	1.13	1.02	1.06
Total number of science courses taken	Mean	2.65	2.73	2.81	2.93
	Standard deviation	0.95	0.94	0.84	0.94
SAT composite score	Mean	814.8	863.2	985.4	949.6
	Standard deviation	150.3	149.2	162.5	158.2

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) by spring 2011. See appendix B for variable definitions.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

b. Data are centered values. Uncentered values are non-Hispanic Black = 94.94, Hispanic = 94.68, non-Hispanic White = 95.12, and other = 96.05.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C5. Descriptive statistics for indicators that predict postsecondary science, technology, engineering, and math success among Texas high school graduates who enrolled in a four-year college, by race/ethnicity

Variable		Non Hispanic Black	Hispanic	Non Hispanic White	Other[a]
High school attendance rate[b] (percent)	Mean	1.27	1.31	1.68	2.67
	Standard deviation	3.16	2.94	2.55	2.09
Grade 10 state assessment scaled score in math	Mean	0.03	0.48	0.92	1.28
	Standard deviation	0.81	0.89	0.92	0.99
Grade 11 state assessment scaled score in math	Mean	–0.06	0.40	0.83	1.18
	Standard deviation	0.80	0.88	0.92	0.97
Grade 10 state assessment scaled score in science	Mean	–0.01	0.34	0.91	1.01
	Standard deviation	0.80	0.83	0.88	0.93
Grade 11 state assessment scaled score in science	Mean	–0.06	0.27	0.82	0.97
	Standard deviation	0.79	0.84	0.91	0.97
Number of Advanced Placement math courses taken	Mean	0.15	0.29	0.39	0.75
	Standard deviation	0.41	0.55	0.59	0.77
Number of Advanced Placement science courses taken	Mean	0.13	0.24	0.30	0.71
	Standard deviation	0.40	0.54	0.59	0.90
Total number of math courses taken	Mean	3.19	3.42	3.55	3.85
	Standard deviation	0.97	0.98	0.82	0.89
Total number of science courses taken	Mean	3.11	3.25	3.34	3.81
	Standard deviation	0.85	0.88	0.77	1.04
SAT composite score	Mean	919.3	980.1	1,117.2	1,164.6
	Standard deviation	156.6	157.0	165.3	182.2

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a four-year Texas public or private college or university (a four-year college) by spring 2011. See appendix B for variable definitions.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

b. Data are centered values. Uncentered values are non-Hispanic Black = 96.67, Hispanic = 96.70, non-Hispanic White = 97.08, and other = 98.07.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C6. Highest math and science courses taken by Texas high school graduates who enrolled in a two-year college, by race/ethnicity (percent)

Subject and course	Non Hispanic Black	Hispanic	Non Hispanic White	Other[a]
Highest math course taken				
Algebra I	4.6	3.7	2.7	2.6
Math Models	4.4	2.8	2.8	2.3
Geometry	19.9	15.5	15.7	12.5
Algebra II	55.7	55.0	51.5	44.2
Precalculus	13.4	19.3	21.9	29.4
Advanced Placement Statistics	0.6	1.0	1.5	2.4
An Advanced Placement calculus course	1.4	2.9	4.0	6.6
Highest science course taken				
Biology	4.6	3.9	3.1	2.8
Integrated Physics and Chemistry	22.3	15.9	17.2	12.5
Chemistry	53.4	53.8	49.1	44.3
Physics	19.1	25.3	29.0	37.8
An Advanced Placement or International Baccalaureate physics course	0.5	0.8	1.2	2.1

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) by spring 2011. See appendix B for variable definitions. Values for an Advanced Placement or International Baccalaureate biology course and an Advanced Placement or International Baccalaureate chemistry course are not included because of small cell sizes.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C7. Highest math and science courses taken by Texas high school graduates who enrolled in a four-year college, by race/ethnicity (percent)

Subject and course	Non Hispanic Black	Hispanic	Non Hispanic White	Other[a]
Highest math course taken				
Geometry	7.2	4.2	2.5	1.4
Algebra II	44.9	32.1	24.9	11.1
Precalculus	32.4	38.5	38.3	30.7
Advanced Placement Statistics	2.7	4.0	6.4	6.9
An Advanced Placement calculus course	10.5	20.4	27.4	49.7
Highest science course taken				
Integrated Physics and Chemistry	7.0	3.1	2.2	1.0
Chemistry	50.2	38.8	33.0	17.6
An Advanced Placement or International Baccalaureate chemistry course	1.2	1.6	1.8	2.7
Physics	37.7	49.4	54.2	60.7
An Advanced Placement or International Baccalaureate physics course	2.8	6.3	8.5	17.8

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a four-year Texas public or private college or university (a four-year college) by spring 2011. See appendix B for variable definitions. Values for Algebra I, Biology, and an Advanced Placement or International Baccalaureate biology course are not included because of small cell sizes.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C8. Postsecondary outcomes among Texas high school graduates who enrolled in a two-year college, by race/ethnicity (percent)

Postsecondary outcome	Non Hispanic Black	Hispanic	Non Hispanic White	Other[a]
Declare a STEM major	10.2	11.7	12.6	14.9
Persist in a STEM major[b,c]	34.9	43.6	44.2	53.8
Complete a degree[c]	6.5	10.3	15.7	17.2
Complete a STEM degree[c]	0.5	1.1	1.7	2.2

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a two-year Texas public college (a two-year college) by spring 2011. See appendix B for outcome definitions.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

b. Analysis includes only college students who declared a STEM major.

c. Data were not available for cohort 7 (which entered grade 9 in 2006).

Source: Authors' analysis of student-level data from the Texas Education Agency.

Table C9. Postsecondary outcomes among Texas high school graduates who enrolled in a four-year college, by race/ethnicity (percent)

Post-secondary outcome	Non Hispanic Black	Hispanic	Non Hispanic White	Other[a]
Declare a STEM major	23.1	28.4	28.8	46.9
Persist in a STEM major[b,c]	43.8	53.3	56.4	67.3
Complete a degree[d]	28.4	37.0	56.7	58.8
Complete a STEM degree[d]	3.5	6.6	9.5	21.3

Note: Data cover students who entered grade 9 in 2000–06, were enrolled in a Texas public high school for at least three years, and enrolled in a four-year Texas public or private college or university (a four-year college) by spring 2011. See appendix B for outcome definitions.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

b. Analysis includes only college students who declared a STEM major.

c. Data were not available for cohort 7 (which entered grade 9 in 2006).

d. Data were not available for cohorts 4–7 (which entered grade 9 in 2003–06).

Source: Authors' analysis of student-level data from the Texas Education Agency.

Appendix D. Output from statistical models

Table D1. Output from model 2, including indicators (students who enrolled in a two-year college)

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Highest math course taken: Algebra I	–0.007*†	0.003	0.007	0.014	–0.003**†	0.001
Highest math course taken: Math Models	–0.005	0.003	–0.041**	0.015	–0.002**	0.001
Highest math course taken: Geometry	–0.003	0.002	–0.017**	0.007	–0.003***†	0.000
Highest math course taken: Precalculus	0.027***†	0.002	0.043***†	0.006	0.014***†	0.001
Highest math course taken: Advanced Placement Statistics	0.014	0.009	0.052*	0.023	–0.002	0.009
Highest math course taken: an Advanced Placement calculus course	0.098***†	0.010	0.117***	0.021	0.037***	0.011
Highest science course taken: Biology	0.005	0.003	0.001	0.013	0.000	0.001
Highest science course taken: an Advanced Placement or International Baccalaureate biology course	0.053	0.039	–0.120	0.107	0.023	0.029
Highest science course taken: Integrated Physics and Chemistry	0.000	0.002	–0.008	0.007	–0.001	0.001
Highest science course taken: an Advanced Placement or International Baccalaureate chemistry course	0.002	0.009	0.000	0.023	–0.003	0.006
Highest science course taken: Physics	0.011***†	0.002	–0.001	0.005	0.002*†	0.001
Highest science course taken: an Advanced Placement or International Baccalaureate physics course	0.033***	0.008	–0.010	0.016	–0.001	0.006
Total number of math courses taken	0.000	0.001	0.006	0.003	0.000	0.000
Number of Advanced Placement math courses taken	0.002	0.009	–0.009	0.017	0.014	0.010
Total number of science courses taken	0.004***	0.001	0.012***†	0.003	0.002***	0.000
Number of Advanced Placement science courses taken	0.025***†	0.003	0.014	0.008	0.014***	0.003
Grade 10 state assessment scaled score in math	0.006***†	0.001	0.005	0.004	0.000	0.001
Grade 11 state assessment scaled score in math	0.014***†	0.001	0.028***†	0.004	0.007***†	0.001
Grade 10 state assessment scaled score in science	0.007***	0.001	0.011**	0.004	–0.001	0.001
Grade 11 state assessment scaled score in science	0.012***†	0.001	0.014***†	0.004	0.004***†	0.001
High school attendance rate	0.001***†	0.000	0.002***	0.001	0.001***†	0.000
Received a high school diploma	–0.009**	0.003	0.027†	0.013	–0.002**	0.001
Number of years enrolled in high school	–0.014***†	0.001	0.013	0.007	–0.004***†	0.000
Home language: not English or Spanish	0.021***†	0.004	0.066***†	0.017	0.013***†	0.003
Home language: Spanish	0.001	0.002	0.029***†	0.007	0.003**†	0.001
Ever an English learner student	0.001	0.004	–0.007	0.019	0.000	0.002
Ever eligible for the federal school lunch program	–0.001	0.001	–0.017***†	0.005	–0.002***†	0.001
Ever in an English as a second language program	0.017***†	0.005	0.035	0.021	0.001	0.002
Ever in special education	0.008***†	0.002	–0.008	0.007	0.000	0.001
Hispanic	0.002	0.001	0.017**†	0.006	–0.001	0.001
Non-Hispanic Black	0.011***†	0.002	–0.043***†	0.008	–0.003***†	0.001
Other race/ethnicity[a]	0.019***†	0.004	0.050***†	0.013	0.003	0.002
Female	–0.109***†	0.002	–0.036***†	0.005	–0.013***†	0.001

(continued)

Table D1. Output from model 2, including indicators (students who enrolled in a two-year college) *(continued)*

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Percentage of students who are English learner students	–0.002***†	0.000	0.000	0.001	0.000**†	0.000
Percentage of students who are eligible for the federal school lunch program	0.000**†	0.000	–0.001***†	0.000	0.000**†	0.000
Percentage of students who are non-Hispanic White	0.000*	0.000	0.000	0.000	0.000	0.000
Percentage of students who are non-Hispanic Black	0.000**†	0.000	0.000	0.000	0.000	0.000
Percentage of students who are "other" race/ethnicity[a]	–0.002***†	0.000	0.001*	0.001	0.000**†	0.000

* Statistically significant at $p < .05$; ** statistically significant at $p < .01$; *** statistically significant at $p < .001$. † Robust to various model specifications.

STEM is science, technology, engineering, and math.

Note: See appendix B for variable definitions and a description of the analytic methodology.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

Table D2. Output from model 2, including indicators (students who enrolled in a four-year college)

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Highest math course taken: Algebra I	0.042***†	0.013	–0.001	0.046	0.025***	0.004
Highest math course taken: Math Models	–0.012	0.012	–0.094	0.053	0.021***	0.004
Highest math course taken: Geometry	0.007	0.005	0.011	0.015	0.011***	0.002
Highest math course taken: Precalculus	0.033***	0.002	0.029***	0.006	–0.005**	0.002
Highest math course taken: Advanced Placement Statistics	–0.033***	0.007	0.019	0.011	–0.060***†	0.007
Highest math course taken: an Advanced Placement calculus course	0.118***	0.008	0.092***†	0.009	0.021**	0.008
Highest science course taken: Biology	0.062***†	0.012	0.018	0.036	0.037***†	0.006
Highest science course taken: an Advanced Placement or International Baccalaureate biology course	0.081	0.057	0.081	0.102	0.094	0.060
Highest science course taken: Integrated Physics and Chemistry	0.037***	0.006	–0.009	0.016	0.023***	0.002
Highest science course taken: an Advanced Placement or International Baccalaureate chemistry course	0.025**	0.009	0.016	0.011	–0.018*	0.007
Highest science course taken: Physics	0.022***	0.003	0.009	0.005	0.001	0.002
Highest science course taken: an Advanced Placement or International Baccalaureate physics course	0.082***	0.007	0.047***†	0.007	0.038***	0.006
Total number of math courses taken	0.015***	0.002	0.011***†	0.003	0.006***	0.001
Number of Advanced Placement math courses taken	0.016**	0.006	0.007	0.005	0.040***†	0.006
Total number of science courses taken	0.038***	0.003	0.018***†	0.003	0.017***	0.002

(continued)

Table D2. Output from model 2, including indicators (students who enrolled in a four-year college) *(continued)*

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Number of Advanced Placement science courses taken	0.044***	0.004	0.004	0.004	0.038***	0.004
Grade 10 state assessment scaled score in math	0.013***	0.002	0.023***	0.003	0.013***	0.001
Grade 11 state assessment scaled score in math	0.028***	0.002	0.036***	0.003	0.015***	0.001
Grade 10 state assessment scaled score in science	0.022***	0.002	0.011***†	0.003	0.001	0.001
Grade 11 state assessment scaled score in science	0.033***	0.002	0.022***†	0.003	0.019***	0.001
High school attendance rate	0.001**	0.000	0.012***†	0.001	0.005***	0.000
Received a high school diploma	0.000	0.010	0.040	0.033	–0.009	0.006
Number of years enrolled in high school	–0.070***†	0.005	–0.042***†	0.010	–0.026***†	0.004
First SAT composite score	0.000**	0.000	0.000	0.000	0.000***	0.000
Home language: not English or Spanish	0.021**†	0.006	0.019*†	0.009	0.025***†	0.007
Home language: Spanish	0.003	0.004	0.034***†	0.007	0.008*†	0.003
Ever an English learner student	0.020	0.013	0.054	0.027	0.023*	0.011
Ever eligible for the federal school lunch program	0.018***†	0.002	–0.002	0.005	0.001	0.002
Ever in an English as a second language program	0.058***†	0.015	0.032	0.031	0.008	0.013
Ever in special education	0.043***†	0.006	0.035**†	0.013	0.016***†	0.003
Hispanic	0.041***†	0.003	0.026***†	0.006	0.008***	0.002
Non-Hispanic Black	0.075***†	0.004	0.040***†	0.007	0.021***†	0.002
Other race/ethnicity[a]	0.071***†	0.006	0.021**†	0.007	0.029***†	0.005
Female	–0.116***†	0.002	–0.048***†	0.003	–0.018***†	0.001
Percentage of students who are English learner students	–0.001**†	0.000	0.000	0.001	0.000	0.000
Percentage of students who are eligible for the federal school lunch program	0.000	0.000	–0.002***†	0.000	0.000***†	0.000
Percentage of students who are non-Hispanic White	–0.001***†	0.000	–0.002***†	0.000	0.000***†	0.000
Percentage of students who are non-Hispanic Black	–0.001***†	0.000	–0.001***†	0.000	0.000***†	0.000
Percentage of students who are "other" race/ethnicity[a]	–0.001***†	0.000	0.001	0.000	0.000	0.000

* Statistically significant at $p < .05$; ** statistically significant at $p < .01$; *** statistically significant at $p < .001$. † Robust to various model specifications.

STEM is science, technology, engineering, and math.

Note: See appendix B for variable definitions and a description of the analytic methodology.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

Table D3. Output from model 4, including indicators and their interactions with race/ethnicity (students who enrolled in a two-year college)

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Highest math course taken: Algebra I	–0.005	0.004	0.016	0.020	–0.003*	0.001
Hispanic * highest math course taken: Algebra I	–0.004	0.006	–0.029	0.026	0.000	0.002
Non-Hispanic Black * highest math course taken: Algebra I	0.004	0.007	0.021	0.035	0.004*	0.002
Other race/ethnicity[a] * highest math course taken: Algebra I	–0.030*	0.015	0.080	0.070	–0.004	0.006
Highest math course taken: Math Models	–0.003	0.004	–0.039	0.022	–0.002	0.001
Hispanic * highest math course taken: Math Models	0.001	0.006	–0.023	0.031	0.000	0.002
Non-Hispanic Black * highest math course taken: Math Models	–0.012	0.007	0.043	0.037	0.000	0.002
Other race/ethnicity[a] * highest math course taken: Math Models	–0.025	0.017	0.037	0.106	0.003	0.006
Highest math course taken: Geometry	0.000	0.002	–0.014	0.010	–0.003***	0.001
Hispanic * highest math course taken: Geometry	–0.005	0.003	–0.010	0.013	0.000	0.001
Non-Hispanic Black * highest math course taken: Geometry	–0.009*	0.004	0.018	0.020	0.002	0.001
Other race/ethnicity[a] * highest math course taken: Geometry	–0.005	0.009	–0.052	0.040	0.000	0.003
Highest math course taken: Precalculus	0.029***	0.002	0.034***	0.008	0.014***	0.001
Hispanic * highest math course taken: Precalculus	–0.007*	0.003	0.020	0.013	–0.001	0.002
Non-Hispanic Black * highest math course taken: Precalculus	0.003	0.005	–0.003	0.017	–0.004	0.002
Other race/ethnicity[a] * highest math course taken: Precalculus	0.015	0.008	0.027	0.029	0.010	0.005
Highest math course taken: Advanced Placement Statistics	0.018	0.013	0.085*	0.035	0.011	0.012
Hispanic * highest math course taken: Advanced Placement Statistics	–0.008	0.020	–0.087	0.052	–0.039*	0.018
Non-Hispanic Black * highest math course taken: Advanced Placement Statistics	–0.022	0.035	–0.052	0.092	0.002	0.025
Other race/ethnicity[a] * highest math course taken: Advanced Placement Statistics	0.034	0.034	–0.031	0.092	0.003	0.038
Highest math course taken: an Advanced Placement calculus course	0.120***	0.014	0.138***	0.031	0.052***	0.014
Hispanic * highest math course taken: an Advanced Placement calculus course	–0.042*	0.020	–0.045	0.046	–0.034	0.021
Non-Hispanic Black * highest math course taken: an Advanced Placement calculus course	–0.105**	0.033	–0.087	0.085	–0.041	0.028
Other race/ethnicity[a] * highest math course taken: an Advanced Placement calculus course	–0.029	0.037	–0.002	0.079	–0.012	0.041
Highest science course taken: Biology	0.016***	0.004	0.015	0.019	0.003	0.002
Hispanic * highest science course taken: Biology	–0.016**	0.005	–0.035	0.025	–0.006**	0.002
Non-Hispanic Black * highest science course taken: Biology	–0.019**	0.007	–0.019	0.035	–0.005*	0.002
Other race/ethnicity[a] * highest science course taken: Biology	–0.024	0.015	0.106	0.076	0.004	0.008

(continued)

Table D3. Output from model 4, including indicators and their interactions with race/ethnicity (students who enrolled in a two-year college) *(continued)*

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Highest science course taken: an Advanced Placement or International Baccalaureate biology course	0.054	0.050	0.106	0.183	0.050	0.052
Hispanic * highest science course taken: an Advanced Placement or International Baccalaureate biology course	0.025	0.074	–0.349	0.212	–0.079	0.052
Non-Hispanic Black * highest science course taken: an Advanced Placement or International Baccalaureate biology course	–0.008	0.096	–0.245	0.362	0.062	0.127
Other race/ethnicity[a] * highest science course taken: an Advanced Placement or International Baccalaureate biology course	–0.298***	0.053	—	—	–0.083	0.055
Highest science course taken: Integrated Physics and Chemistry	0.004	0.003	0.008	0.010	0.001	0.001
Hispanic * highest science course taken: Integrated Physics and Chemistry	–0.007*	0.003	–0.022	0.015	–0.004***	0.001
Non-Hispanic Black * highest science course taken: Integrated Physics and Chemistry	–0.007	0.004	–0.059**	0.020	–0.003**	0.001
Other race/ethnicity[a] * highest science course taken: Integrated Physics and Chemistry	–0.002	0.010	–0.042	0.041	–0.004	0.004
Highest science course taken: an Advanced Placement or International Baccalaureate chemistry course	0.004	0.011	–0.020	0.031	0.002	0.008
Hispanic * highest science course taken: an Advanced Placement or International Baccalaureate chemistry course	–0.001	0.019	0.024	0.053	–0.020	0.012
Non-Hispanic Black * highest science course taken: an Advanced Placement or International Baccalaureate chemistry course	–0.017	0.031	0.147	0.135	0.005	0.018
Other race/ethnicity[a] * highest science course taken: an Advanced Placement or International Baccalaureate chemistry course	0.003	0.045	0.169	0.110	0.008	0.036
Highest science course taken: Physics	0.009**	0.003	–0.001	0.008	0.001	0.001
Hispanic * highest science course taken: Physics	0.007*	0.003	–0.010	0.011	0.003	0.002
Non-Hispanic Black * highest science course taken: Physics	–0.001	0.004	0.024	0.017	0.002	0.002
Other race/ethnicity[a] * highest science course taken: Physics	0.005	0.007	0.044	0.026	0.005	0.005
Highest science course taken: an Advanced Placement or International Baccalaureate physics course	0.023*	0.011	–0.024	0.021	–0.006	0.008
Hispanic * highest science course taken: an Advanced Placement or International Baccalaureate physics course	0.039*	0.016	0.027	0.036	0.004	0.012
Non-Hispanic Black * highest science course taken: an Advanced Placement or International Baccalaureate physics course	–0.041	0.021	0.136	0.073	0.024	0.017
Other race/ethnicity[a] * highest science course taken: an Advanced Placement or International Baccalaureate physics course	0.014	0.030	0.013	0.059	0.035	0.024

(continued)

Table D3. Output from model 4, including indicators and their interactions with race/ethnicity (students who enrolled in a two-year college) *(continued)*

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Total number of math courses taken	0.001	0.001	0.005	0.004	0.001	0.001
Hispanic * total number of math courses taken	–0.001	0.002	0.000	0.006	–0.001	0.001
Non-Hispanic Black * total number of math courses taken	–0.001	0.002	0.007	0.009	–0.001	0.001
Other race/ethnicity[a] * total number of math courses taken	–0.007	0.004	0.006	0.014	–0.001	0.002
Number of Advanced Placement math courses taken	–0.010	0.012	–0.029	0.027	0.004	0.013
Hispanic * number of Advanced Placement math courses taken	0.017	0.017	0.038	0.039	0.023	0.018
Non-Hispanic Black * number of Advanced Placement math courses taken	0.071*	0.030	0.075	0.070	0.006	0.024
Other race/ethnicity[a] * number of Advanced Placement math courses taken	0.021	0.028	0.007	0.062	0.016	0.035
Total number of science courses taken	0.009***	0.002	0.014**	0.005	0.003***	0.001
Hispanic * total number of science courses taken	–0.007***†	0.002	–0.003	0.007	–0.003***	0.001
Non-Hispanic Black * total number of science courses taken	–0.011***	0.002	–0.007	0.011	–0.003**	0.001
Other race/ethnicity[a] * total number of science courses taken	–0.008	0.005	–0.005	0.017	–0.002	0.003
Number of Advanced Placement science courses taken	0.021***	0.004	0.017	0.011	0.013***	0.003
Hispanic * number of Advanced Placement science courses taken	0.003	0.006	–0.018	0.018	0.002	0.005
Non-Hispanic Black * number of Advanced Placement science courses taken	0.002	0.010	0.004	0.036	–0.011	0.006
Other race/ethnicity[a] * number of Advanced Placement science courses taken	0.010	0.013	0.015	0.030	–0.009	0.012
Grade 10 state assessment scaled score in math	0.005***	0.001	0.006	0.005	0.000	0.001
Hispanic * grade 10 state assessment scaled score in math	0.002	0.002	0.003	0.008	0.000	0.001
Non-Hispanic Black * grade 10 state assessment scaled score in math	0.000	0.002	–0.003	0.011	–0.001	0.001
Other race/ethnicity[a] * grade 10 state assessment scaled score in math	0.001	0.005	–0.017	0.018	0.000	0.003
Grade 11 state assessment scaled score in math	0.014***	0.001	0.028***	0.005	0.008***	0.001
Hispanic * grade 11 state assessment scaled score in math	0.000	0.002	0.004	0.007	0.000	0.001
Non-Hispanic Black * grade 11 state assessment scaled score in math	–0.005*	0.002	–0.012	0.010	–0.004***†	0.001
Other race/ethnicity[a] * grade 11 state assessment scaled score in math	0.010	0.005	0.001	0.017	0.000	0.003
Grade 10 state assessment scaled score in science	0.009***	0.001	0.019***	0.005	0.000	0.001
Hispanic * grade 10 state assessment scaled score in science	–0.003	0.002	–0.014	0.008	0.000	0.001
Non-Hispanic Black * grade 10 state assessment scaled score in science	–0.008***	0.002	–0.016	0.011	–0.002*	0.001
Other race/ethnicity[a] * grade 10 state assessment scaled score in science	–0.005	0.005	–0.024	0.018	0.001	0.003

(continued)

Table D3. Output from model 4, including indicators and their interactions with race/ethnicity (students who enrolled in a two-year college) *(continued)*

Variable	Declare a STEM major		Persist in a STEM major		Complete a STEM degree	
	Beta	Standard error	Beta	Standard error	Beta	Standard error
Grade 11 state assessment scaled score in science	0.015***	0.001	0.008	0.005	0.004***	0.001
Hispanic * grade 11 state assessment scaled score in science	–0.005**†	0.002	0.013	0.007	0.000	0.001
Non-Hispanic Black * grade 11 state assessment scaled score in science	–0.006**	0.002	0.006	0.010	–0.001	0.001
Other race/ethnicity[a] * grade 11 state assessment scaled score in science	–0.006	0.005	0.018	0.016	0.008*	0.003
High school attendance rate	0.000	0.000	0.003***	0.001	0.001***	0.000

* Statistically significant at $p < .05$; ** statistically significant at $p < .01$; *** statistically significant at $p < .001$. † Robust to various model specifications.

— is not available because of an insufficient number of observations. STEM is science, engineering, technology, and math.

Note: See appendix B for variable definitions and a description of the analytic methodology.

a. Refers to non-Hispanic American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander students.

Source: Authors' analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

Appendix E. Tests of joint significance of indicators

For each regression model, Wald tests were used to evaluate whether terms in the model were jointly equal to zero (table E1). The Wald test for the joint significance of the math and science variables tests whether all indicators are jointly equal to zero. The Wald test for the Hispanic variables tests whether all interaction terms between Hispanic status and indicators are jointly equal to zero. Rejecting the null hypothesis indicates that the terms being evaluated in the Wald test are not jointly equal to zero and therefore as a block are considered to improve the predictiveness of the model.

Table E1. Summary statistics for Wald tests of the variable blocks

	Students who enrolled in a two year college			Students who enrolled in a four year college		
Variable block	Declare a STEM major	Persist in a STEM major	Earn a STEM degree	Declare a STEM major	Persist in a STEM major	Earn a STEM degree
Math and science	137.29***	57.24***	57.35***	574.94***	267.33***	202.11***
Hispanic	5.79***	1.72***	5.39***	7.53***	5.33***	10.55***

*** Statistically significant at $p < .001$.

Note: Data in table are *F*-statistics.

Source: Author's analysis of student-level data from the Texas Education Agency and the Texas Higher Education Coordinating Board.

Notes

1. For example, if taking advanced math and science courses in high school is an indicator of postsecondary STEM outcomes, future studies could investigate whether variation in the availability of such courses accounts for variations in postsecondary STEM outcomes.
2. Wald tests indicated that the indicator terms in the regression model were not jointly equal to zero and therefore as a block are considered to improve the predictiveness of the model (see appendix E).
3. The analysis of differential functioning of indicators is limited to indicators that were both significant and robust predictors of postsecondary STEM success.
4. The inclusion of the interaction terms improved the predictiveness of statistical models for both students who enrolled in a two-year college and students who enrolled in a four-year college (see table E1 for appendix E for the summary statistics for Wald tests of joint significance).
5. Relationships between highest science course taken and postsecondary STEM success follow a similar pattern, reflecting the fact that Physics is one level higher than the typical highest course for students who enroll in a two-year college and Advanced Placement Physics is one level higher than the typical highest course for students who enroll in a four-year college.
6. The maximum possible delay between high school graduation and college matriculation varied by cohort (see table B1 in appendix B).
7. Because these analyses are ancillary to the main purpose of the study, the study team did not conduct robustness checks. This section highlights interaction terms that exceed the significance level of $p < .01$.
8. Students in later cohorts had fewer semesters of available data in which to observe their leaving a STEM major, so by an artifact of this data availability, these cohorts appear to have higher rates of persistence.
9. The groups were analyzed separately because sub-bachelor's and bachelor's programs prepare students to enter different types of occupations (Carnevale, Smith, & Melton, 2011; Rothwell, 2013). For example, within engineering and engineering technologies, associate's level STEM degree programs, offered by two-year colleges, prepare students to be technicians, whereas bachelor's level STEM degree programs, offered by four-year colleges, prepare students to be engineers (see Schneider, 2016, for a summary of median wages for Texas college graduates by major and degree type).
10. Although this document refers to a school year outside the scope of the study, the course names referenced therein matched the courses names in the dataset.
11. Statistics was taught only at the Advanced Placement level in Texas during the period under examination.
12. Calculus was taught only at the Advanced Placement level in Texas during the period under examination.
13. The use of ordinary least squares for binary outcomes is not without its detractors. Wooldridge (2002) and Greene (2003) discuss the strengths and limitations of the ordinary least squares specification, as well as adjustments, which were adopted in the current study, for mitigating some of the potential biases that result from this estimation approach. Wooldridge (2002) and Greene (2003) also report on the conditions under which the linear probability model provides reasonable estimates of the average marginal effects across values of the covariates of interest. Horrace and Oaxaca (2006) formally demonstrate the potential bias and inconsistency of the estimators derived

from the linear probability model. Nonetheless, the technique has been used widely in applied econometric work, including by Heckman and Snyder (1997) and Angrist (2001).

14. Although only indicators identified in Hinojosa et al. (2016) are described in this section, one exploratory indicator was also examined by the model (see table B2).
15. That is, indicators that are in each subsequent model and so, here, the reduced model is nested within the full model.

References

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago Public High Schools: A close look at course grades, failures, and attendance in the freshman year*. Chicago. IL: Consortium on Chicago School Research. http://eric.ed.gov/?id=ED498350

Angrist, J. D. (2001). Estimations of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice: Reply. *Journal of Business & Economic Statistics, 19*(1), 2–16.

Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching, 47*(5), 564–582. http://eric.ed.gov/?id=EJ887585

Beede, D., Julian, T., Khan, B., Lehrman, R., McKittrick, G., Langdon, D., & Doms, M. (2011). *Education supports racial and ethnic equality in STEM* (ESA Issue Brief No. 05–11). Washington, DC: U.S. Department of Commerce. http://eric.ed.gov/?id=ED523768

Burge, S. W. (2013). Cohort changes in the relationship between adolescents' family attitudes, STEM intentions and attainment. *Sociological Perspectives, 56*(1), 49–73.

Carnevale, A. P., Smith, N., & Melton, M. (2011). *STEM: Science technology engineering mathematics*. Washington, DC: Georgetown University Center on Education and the Workforce. Retrieved May 19, 2017, from https://cew.georgetown.edu/wp-content/uploads/2014/11/stem-complete.pdf.

DeWitt, J., Archer, L., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2011). High aspirations but low progression: The science aspirations–careers paradox amongst minority ethnic students. *International Journal of Science and Mathematics Education, 9*(2), 243–271. http://eric.ed.gov/?id=EJ918490

Foley, E., Mishook, J., Thompson, J., Kubiak, M., Supovitz, J., & Rhude-Faust, M. K. (2008). *Beyond test scores: Leading indicators for education.* Providence, RI: Annenberg Institute for School Reform at Brown University. http://eric.ed.gov/?id=ED533117

Ginsburg, A., Jordan, P., & Chang, H. (2014). *Absences add up: How school attendance influences student success*. Washington, DC: Attendance Works. Retrieved October 10, 2016, from http://www.attendanceworks.org/wordpress/wp-content/uploads/2014/09/Absenses-Add-Up_September-3rd-2014.pdf.

Greene, W. H. (2003). *Econometric analysis*. Upper Saddle River, NJ: Prentice Hall.

Griffith, A. 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. http://eric.ed.gov/?id=EJ905602

Heckman, J. J., & Snyder Jr., J. M. (1997). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators (Special Issue). *RAND Journal of Economics, 28*(0): S142–S189.

Higher Education Research Institute. (2010). *Degrees of success: Bachelor's degree completion rates among initial STEM majors*. HERI Research Brief. Los Angeles, CA: Author. Retrieved October 13, 2016, from http://www.heri.ucla.edu/nih/downloads/2010%20-%20Hurtado,%20Eagan,%20Chang%20-%20Degrees%20of%20Success.pdf.

Hinojosa, T., Rapaport, A., Jaciw, A., LiCalsi, C., & Zacamy, J. (2016). *Exploring the foundations of the future STEM workforce: K–12 indicators of postsecondary STEM success* (REL 2016–122). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. http://eric.ed.gov/?id=ED565641

Horrace, W. C., & Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters, 90*(3), 321–327. Retrieved October 11, 2016, from http://surface.syr.edu/cgi/viewcontent.cgi?article=1138&context=ecn.

Kokkelenberg, E. C., & Sinha, E. (2010). Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students. *Economics of Education Review, 29*(6), 935–946. http://eric.ed.gov/?id=EJ905609

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). *STEM: Good jobs now and for the future* (ESA Issue Brief No. 03–11). Washington, DC: U.S. Department of Commerce. http://eric.ed.gov/?id=ED522129

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken, NJ: Wiley.

National Center for Education Statistics. (2011). *Postsecondary awards in science, technology, engineering, and mathematics, by state: 2001 and 2009* (NCES No. 2011–226). National Center for Education Statistics Working Paper. Washington, DC: U.S. Department of Education. http://eric.ed.gov/?id=ED518801

National Center for Education Statistics. (2016). *Digest of education statistics: 2015* (NCES No. 2016–014, Tables 318.45 and 322.20). Washington, DC: U.S. Department of Education. Retrieved February 3, 2017, from https://nces.ed.gov/programs/digest/d15/.

Neild, R. C., Balfanz, R., & Herzog, L. (2007). An early warning system. *Educational Leadership, 65*(2), 28–33.

Riegle-Crumb, C., & King, B. (2010). Questioning a White male advantage in STEM: Examining disparities in college major by gender and race/ethnicity. *Educational Researcher, 39*(9), 656–664. http://eric.ed.gov/?id=EJ912229

Rothwell, J. (2013). *The hidden STEM economy.* Washington, DC: Brookings Institution. Retrieved May 19, 2017, from https://www.brookings.edu/research/the-hidden-stem-economy/.

Ryan, S., & Ream, R. K. (2016). Variation across Hispanic immigrant generations in parent social capital, college-aligned actions, and four-year college enrollment. *American Educational Research Journal, 53*(4), 953–986. http://eric.ed.gov/?id=EJ1111139

Schneider, M. (2016). *Majors matter: Differences in wages over time in Texas*. Washington, DC: College Measures and American Institutes for Research. Retrieved May 19, 2017, from http://tx.edpays.org/Content/MajorsMatter_TX_EMS.pdf.

Shaw, E. J., & Barbuti, S. (2010). Patterns of persistence in intended college major with a focus on STEM majors. *NACADA Journal, 30*(2), 19–34. http://eric.ed.gov/?id=EJ906475

StataCorp. (2013). *Stata statistical software: Release 13*. College Station, TX: Author.

Texas Education Agency. (2012). *Graduation credit requirements for students entering ninth grade beginning in 2012–2013*. Austin, TX: Author. Retrieved August 3, 2016, from http://www.tea.state.tx.us/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=2147509385&libID=2147509371.

Texas Education Agency. (2016). *Snapshot 2015: State totals.* Austin, TX: Author. Retrieved October 11, 2016, from https://rptsvr1.tea.texas.gov/perfreport/snapshot/2015/state.html.

Texas Guaranteed Student Loan Corporation. (2013). *State of student aid and higher education in Texas*. Round Rock, TX: Author.

Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk, 12*(3), 243–270. http://eric.ed.gov/?id=EJ780680

Vilorio, D. (2014). STEM 101: Intro to tomorrow's jobs. *Occupational Outlook Quarterly, Spring.* Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics. Retrieved October 11, 2016, from http://www.bls.gov/careeroutlook/2014/spring/art01.pdf.

Wang, X. (2013a). Modeling entrance into STEM fields of study among students beginning at community colleges and four-year institutions. *Research in Higher Education, 54*(6), 664–692. http://eric.ed.gov/?id=EJ1039145

Wang, X. (2013b). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal, 50*(5), 1081–1121. http://eric.ed.gov/?id=EJ1018916

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

You, S. (2013). Gender and ethnic differences in precollege mathematics coursework related to science, technology, engineering, and mathematics (STEM) pathways. *School Effectiveness and School Improvement, 24*(1), 64–86. http://eric.ed.gov/?id=EJ995084

The Regional Educational Laboratory Program produces 7 types of reports

Making Connections
Studies of correlational relationships

Making an Impact
Studies of cause and effect

What's Happening
Descriptions of policies, programs, implementation status, or data trends

What's Known
Summaries of previous research

Stated Briefly
Summaries of research findings for specific audiences

Applied Research Methods
Research methods for educational settings

Tools
Help for planning, gathering, analyzing, or reporting data or research