Does Participation in Advanced High School Math Courses Predict Postsecondary Enrollment?

Erin R. Twomey-Wilson[1]

[1]The College Crusade of Rhode Island, 134 Thurbers Avenue, Providence, RI 02905

Abstract

The College Crusade of Rhode Island's mission is to prepare and inspire young people in the state to become the first in their families to attend and complete college. As Rhode Island's most comprehensive college access and success program, the relationship between advanced math courses and college enrollment is a major focus area of our work. We hypothesize that program participants who complete more advanced math courses during their final year of high school will have higher probabilities of enrolling in college. Contingency table analysis was conducted using logistic regression, estimated odds, estimated probability, Pearson chi-squared statistic, and the likelihood-ratio statistic. We observed that program participants who take advanced math courses in high school have a higher probability of enrolling in college. These results provide useful information in ongoing efforts to increase the numbers of low-income and first-generation students who successfully enroll in college.

Key Words: Categorical Data Analysis, Contingency Table Analysis, Logistic Regression, Estimated Odds, Chi-squared Tests of Independence, Estimated Probability

1. Introduction

The College Crusade, founded in 1989, is Rhode Island's most comprehensive college access program. Our mission is to inspire and prepare young people in Rhode Island to become the first in their family to attend and complete college.

Every year, we provide college readiness programs and personalized advisory services to over 4,000 middle school, high school and college students from Providence, Pawtucket, Central Falls, Woonsocket, Cranston and West Warwick. All our students are low-income (based on free/reduced price lunch status at the time of enrollment), 95% identify as students of color, and approximately 93% are first-generation students.

The College Crusade recruits a new class of sixth grade students every year and provides them with ongoing services through their first year of college. These services include personalized support through 29 advisors who operate in 37 middle schools and high schools and at the state's three public colleges. We also provide over 60 innovative programs to meet students' academic, social and emotional development, career education, and college readiness needs.

The College Crusade's program design served as a model in 1997 for the U.S. Department of Education's Gaining Early Awareness and Readiness for Undergraduate Programs (GEAR UP) initiative. As the federal government's largest investment in college access, GEAR UP currently serves over 700,000 students in 44 states. The College Crusade has designed, administered, and implemented Rhode Island GEAR UP program since its inception in 1999, which provides $3.5 million annually and has rigorous program and reporting requirements. A piece of the reporting requirements focuses on the percentages of students who pass the required math courses: Pre-Algebra, and Algebra I. Reporting requirements also include the percentage of students who take 2 years of math beyond Algebra I by 12th grade.

Studies suggest that inequities in access to and taking of advanced math are central to differences in the educational attainment of youth from racial/ethnic minority, low socioeconomic status, and English Language Learners backgrounds. (Byun, et.al, 2015) As Rhode Island's most comprehensive college access and success program, the relationship between advanced math courses and college enrollment is a major focus area of our work.

2. Purpose of Study

The work of Long, et. al. (2012) suggests that requiring or encouraging students to enroll in even one rigorous mathematics course in their first 2 years of high school can substantially improve graduation and 4-year college enrollment rates. The College Crusade of Rhode Island's focus is to increase the numbers of low-income and first-generation students who successfully enroll in college.

The purpose of this study is to examine the relationship between advanced math courses and college enrollment of our program participants. We hypothesize that program participants who complete advanced courses during their last year of high school will have higher probabilities of immediately enrolling in college.

3. Data and Methods

3.1 Student Demographics

The sample consisted of 898 high school program participants that completed (i.e. passed) Algebra II or higher by grade 12. The sample program participants are low-income (based on free/reduced price lunch status at the time of enrollment) and 95% identify as students of color, and approximately 93% are first-generation students. Tables 1 through 3 summarize the sample demographics by sex, race/ethnicity, and home language.

Table 1: Breakdown of program participants by sex. Sex is defined as the gender on the program participant's birth certificate at time of program enrollment.

Sex	Number of Participants	Percentage
Female	521	58%
Male	377	42%
Grand Total	898	100%

Table 2: Breakdown of program participants by race/ethnicity. Program participants who identified as Hispanic or Latino were removed from the race categories in order not to be counted twice in the analysis.

Race/Ethnicity	Number of Participants	Percentage
Hispanic or Latino	504	56%
Black or African American	184	20%
Two or More Races	117	13%
White	48	5%
Asian	27	3%
Unknown	14	2%
American Indian or Alaska Native	2	0%
Native Hawaiian or Other Pacific Islander	2	0%
Grand Total	898	100%

Table 3: Breakdown of program participants by home language. Home language is defined as the primary language spoke in the home.

Home Language	Number of Students	Percentage
English	457	51%
Spanish	369	41%
Unknown	53	6%
Portuguese	7	1%
Cape Verdean, Creole	4	0%
Other	2	0%
French	2	0%
Hmong	2	0%
Laotian	1	0%
Cambodian	1	0%
Grand Total	898	100%

3.2 Course Classification

The Rhode Island GEAR UP program requires that the College Crusade track and measure student completion of core mathematics courses. These measurements include: 1. The number of students that pass Pre-algebra by the end of eighth grade, 2. The number of students that pass Algebra I by the end of ninth grade, and 3. The number of students that take 2 years of mathematics courses beyond Algebra I by the twelfth grade. The third objective is the center of this study and includes the following courses: Algebra II, Advanced Math, Trigonometry, Statistics, Pre-calculus, and Calculus.

Following the construction of the mathematics course credit measures of the National Center for Education Statistics (NCES), the National Assessment of Educational Process equivalent mathematics classifications was used to isolate appropriate courses to include in the course-taking measures. The set of six GEAR UP courses was classified into four major subdivisions: 1. Middle Academic II, 2. Advanced I, 3. Advanced II, and 4. Advanced III. (National Center for Education Statistics, 2003)

The four-level NCES subdivision method was applied to the GEAR UP course classifications and descriptions. Middle Academic II courses included Algebra II and the equivalent. Advanced I included Trigonometry, Statistics, and Probability. Advanced II included Pre-calculus, while Advanced III included Calculus. Table 4 is a summary of the relationship between the GEAR UP Classifications and the NCES subdivisions.

Table 4: Breakdown of mathematics by GEAR UP classification and NCES subdivisions.

	NCES Subdivisions				
GEAR UP Course	**Middle Academic II**	**Advanced I**	**Advanced II**	**Advanced III**	**Grand Total**
Advanced Math	169				169
Algebra II	219				219
Calculus				144	144
Pre-Calculus			279		279
Statistics		72			72
Trigonometry		15			15
Grand Total	388	87	279	144	898

The course subdivisions are designed the capture the nature of the highest-level mathematics course completed by grade 12, not attempted. (National Center for Education Statistics, 2003)

3.3 College Enrollment

The National Student Clearinghouse Student Tracker provided records used to verify college enrollment for program participants. StudentTracker is the only nationwide source of college enrollment and degree data. More than 3,700 colleges and universities — enrolling over 99 percent of all students in public and private U.S. institutions — regularly provide enrollment and graduation data to the Clearinghouse. (National Student Clearinghouse, 2020)

Records were utilized to determine if a student enrolled in either a two- or four- year institution. College Crusade program participants sign a waiver form allowing the organization to gain access to enrollment data. If a student had an enrollment record, they were counted as having enrolled in college.

Table 5: Breakdown of mathematics classifications and college enrollment.

	College		
Grouping	**Yes**	**No**	**Total**
Middle Academic II	321	67	388
Advanced I	80	7	87
Advanced II	260	19	279
Advanced III	140	4	144
Total	801	97	898

3.4 Analytical Methods

Exploratory data analysis included using the Pearson's Chi-squared Test and the Log Likelihood Ratio (G-Test) to test for the level of association between college enrollment and classification level. Logistic regression was then used to analyze the effect of the course classification level on college enrollment.

3.4.1 Tests of Independence: Pearson's Chi-Squared and Log Likelihood Ratio

The Chi-Squared test of independence is used to determine if there is an association between two categorical variables. The frequency of each category for one categorical variable is compared across the category of the second categorical variable. The Person chi-squared statistics for testing H_o (X and Y are independent) is:

$$\chi^2 = \sum \frac{\left(n_{ij} - \mu_{ij}\right)^2}{\mu_{ij}} \tag{1}$$

with the sum taken over all the cells of the table and where the expected frequency μ_{ij}= n_{ij} (cell count). The p-value is the probability, under $H_{o,}$ that χ^2 is at least as large of the observed variable. (Agresti, 2019)

An alternative to the Chi-Squared test of independence is the likelihood method for significance tests. The likelihood-ratio test determines the parameter values that maximize the likelihood function (a) under the assumption that H_o (X and Y are independent) is true and (b) under the more general condition that H_o may or may not be true. For two-way contingency tables with likelihood function based on the multinomial distribution, the likelihood ratio chi-squared statistic is:

$$G^2 = 2 \sum n_{ij} log \left(\frac{n_{ij}}{\mu_{ij}}\right) \tag{2}$$

Like the Pearson χ^2 and likelihood ratio G^2 provide separate test statistics, they share similar properties and provide the same conclusions. When H_o (X and Y are independent) is true and the expected frequencies are large, the two statistics share the same chi-squared distribution and similar numbers. They also have limitations as they indicate the degree of evidence for an association and rarely adequate for answering all study questions. (Agresti, 2019)

3.4.2 Logistic Regression

Logistic regression is the most popular model for binary data. The logistic regression model has a linear form for the logit of the success of probability; the logarithm of the odds,

$$logit[\hat{\pi}(x)] = log\left[\frac{\hat{\pi}(x)}{1 - \hat{\pi}(x)}\right] = \alpha + \beta x \tag{3}$$

The corresponding formula for $\pi(x)$, using the exponential function exp $(\alpha + \beta x)$ = $e^{\alpha+\beta x}$,

$$\hat{\pi}(x) = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}. \tag{4}$$

The effect parameter β determines the rate of increased or decrease of the curve for $\pi(x)$. The sign of β indicates whether the curve ascends (β>0) ore descends (β<0). When β=0, the curve flattens to a horizontal straight line and the binary response variable is the independent of the explanatory variable. (Agresti, 2019)

3.4.3 Odds of Success and Odds Ratio

The logistic regression model above indicates that the logit increases by β for every 1-unit increase in x. Exponentiating both sides of the logistic regression model gives an interpretation of the odds and odds ratio. The odds of success are:

$$\frac{\hat{\pi}(x)}{1 - \hat{\pi}(x)} = \exp(\alpha + \beta x) = e^{\alpha}\left(e^{\beta}\right)e^{x} \tag{5}$$

Therefore, the odds multiply by e^{β} for every c-unit increase in x. (Agresti, 2019, Bilder and Loughin, 2015)

To determine how the odds of success have changed by this c-unit increase, the odds ratio is:

$$\frac{Odds_{x+c}}{Odds_{x}} = \frac{\exp(\alpha + \beta_{x+c})}{\exp(\alpha + \beta x)} = \exp(c\beta) \tag{6}$$

The standard interpretation of this odds ratio is the odds of a success change by exp $(c\beta)$ times for every c-unite increase in x. (Bilder and Loughin, 2015)

3.4.4 Statistical Inference: Wald and Likelihood Ratio Tests

Hypothesis tests can be used to assess the importance of explanatory variables in a model. A test of H_o: β = 0 vs H_a: $\beta \neq 0$ evaluates the explanatory term in the logistic regression model above. If β = 0, the corresponding explanatory term is excluded from the model. If $\beta \neq 0$, the corresponding term is included in the model. A Wald test or Likelihood ratio test is used to evaluate β. (Bilder and Loughin, 2015)

The Wald statistic is:

$$Z_o = \frac{\hat{\beta}}{SE} \tag{7}$$

where SE is the unrestricted standard error of $\hat{\beta}$. If the null hypothesis is true, Z_o has an approximate normal standard distribution. The confidence interval for the Wald statistic is:

$$\hat{\beta} \pm z_{\frac{\alpha}{2}}(SE) \tag{8}$$

The p-value is defined as:

$$2P(Z > |Z_o|) \tag{9}$$

where Z has a standard normal distribution. (Agresti, 2019, Bilder and Loughin, 2015)

The Likelihood ratio test statistic equals:

$$2\log\left(\frac{l_1}{l_0}\right) = 2[\log(l_1) - \log(l_0)] = 2(L_1 - L_0) \tag{10}$$

where L_0 and L_1 denote the maximized log-likelihood functions. The maximized value of the likelihood function by l_o under H_o: β=0 and by l_1 when β need not equal zero. The likelihood ratio statistic can be informally written as:

$$\Lambda = \frac{Maximum\ of\ likelihood\ function\ under\ H_o}{Maximum\ of\ likelihood\ function\ under\ H_o\ or\ H_a} \tag{11}$$

(Agresti, 2019, Bilder and Loughin, 2015)

4. Results

It has been shown that advanced math course taking has had positive effects on college enrollment. (Byun, Soo-Yong et al., 2015) Therefore, results described herein focus o the relationship between math course classification level (1 = Middle Academic II, 2 = Advanced 1, 3 = Advanced II, and 4 = Advanced III) and college enrollment. Analysis of the relationship using tests of independence and logistic regression may help determine if college enrollment is dependent upon math course classification level ($\beta \neq 0$) or not (β=0).

4.1 Tests of Independence

4.1.1 Pearson's Chi-squared Test of Independence (χ^2)

A significant association between math course classification level and college enrollment is show by the Chi-squared test of independence. Table 6 is a summary of the Pearson's Chi-Squared test of independence results.

Table 6: A summary of the Pearson's Chi-squared test of independence.

Test of Independence	Test Statistic (χ^2)	Degrees of Freedom	P-value
Pearson's Chi-Squared	31.76	3	5.88×10^{-7}

The test results provide strong evidence that there is an association between college enrollment and math course classification level (P < 0.0001).

4.1.2 Likelihood Ratio Statistic (G^2)

A significant association between math course classification level and college enrollment is shown by the likelihood ratio statistic. Table 7 is a summary of the likelihood ratio statistic results.

Table 7: A summary of the likelihood ratio statistic.

Test of Independence	Test Statistic (G^2)	Degrees of Freedom	P-value
Likelihood Ratio Statistic	33.79	3	2.20 x 10-7

The test results provide strong evidence that there is an association between college enrollment and math course classification level (P < 0.0001).

4.2 Logistic Regression

4.2.1 Logistic Regression Model

The logistic regression model fit is:

$$logit[\hat{\pi}(x)] = log\left[\frac{\hat{\pi}(x)}{1-\hat{\pi}(x)}\right] = 0.9864 + 0.5898x \quad (12)$$

The summary of the logistic regression output is provided in Table 8 below.

Table 8: A summary of the logistic regression model output.

Variable	Estimate	Standard Error	Z-Value	P-Value
Intercept	0.9864	0.2147	4.595	4.34×10^{-6}
Classification	0.5898	0.1125	5.241	1.60×10^{-6}

4.2.2 Estimated Probability

The estimated probability of college enrollment is:

$$\hat{\pi}(x) = \frac{\exp(0.9864 + 0.5898x)}{1 + \exp(0.9864 + 0.5898x)}. \quad (13)$$

Since β = 0.5898 is greater than zero, the estimated probability is higher at higher classification levels. The effect of classification level seems relatively strong in that $\hat{\pi}(x)$

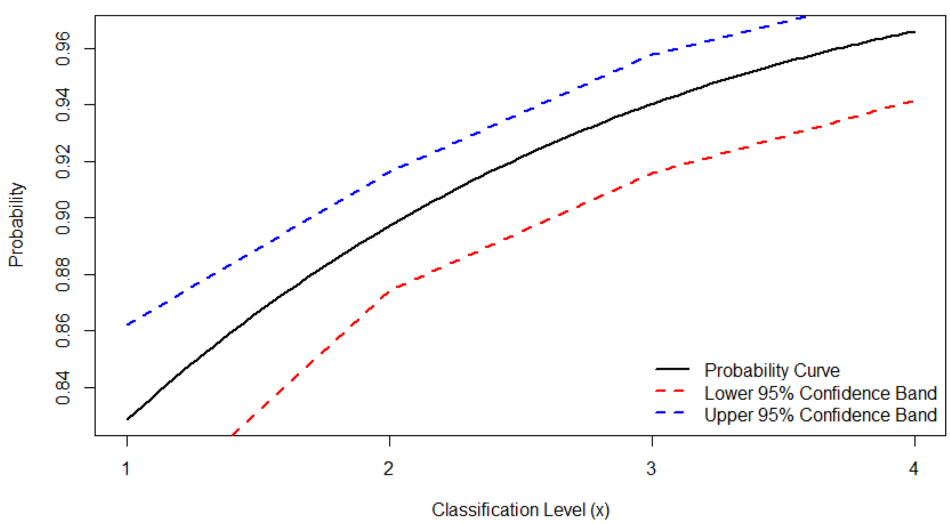
changes over the ranges of x values. The estimated probability of college enrollment can be seen in Table 9 below.

Table 9: Estimated probability of college enrollment based on classification level.

Classification Level (x)	$\hat{\pi}(x)$
Middle Academic II (1)	0.8287
Advanced I (2)	0.8972
Advanced II (3)	0.9402
Advanced III (4)	0.9660

At the maximum classification level of 4, the estimated probability of a program participant enrolling in college equals 0.9660. Classification level has an effect in the sense that $\hat{\pi}(x)$ changes over the range of x values. Figure 1 plots the estimated probabilities.

Figure 1: Linear approximation to the logistic regression curve.



4.2.3 Estimated Odds and Odds Ratios

The logistic regression model shows β = 0.5898. The estimated odds of a program participant enrolling in college multiply by exp(β) = exp(0.5898) = 1.80 for each increase in classification level. In other words, there is an 80% increase in the odds of a program participant enrolling in college with a 1 level increase in classification level. For example, at student that took Pre-Calculus (x=3) their senior year in high school has an 80% increase in odds of enrolling in college than those who took Trigonometry or Statistics (x=2). A summary of the estimated odds and odds ratios can be seen in Table 10 below.

Table 10: Odds of Success and Odds Ratios at the 95% Confidence level generated by the Logistic Regression Model where x is the course classification level: 1 = Middle Academic II, 2 = Advanced 1, 3 = Advanced II, and 4 = Advanced III.

x	Estimated Odds	Odds Ratio	Lower C.I.	Upper C.I.
1	4.8365	1.0000	1.0000	1.0000
2	8.7233	1.8036	1.4579	2.2699
3	15.7336	3.2531	2.1254	5.1526
4	28.3776	5.8673	3.0986	11.6959

The estimated odds ratio of a program participant enrolling in college by a one level in crease in classification is defined by exp(cβ). For example, the odds ratio between classification level x = 1 and x = 3 is exp(2*0.5898) = 3.2531. The estimated odds of program participants enrolling in college completing an x = 3 classification course are 3.25 times the odds the program participants who complete an x=1 classification course.

4.3 Statistical Inference

4.3.1 Wald Test

With $\beta = 0.5898$ and the standard error (SE) = 0.1125, the 95% Wald confidence interval for β is (0.3698,0.8197). The Wald statistic, $z = \beta/SE = 0.5898/0.1125=5.2427$ with p-value < 0.0001 shows strong evidence of a positive effect of classification level on college enrollment. The equivalent of the Chi-square statistic is $z^2 = 5.2427^2 = 27.4859$, df =1, p value < 0.0001. A summary of the Wald test can be seen in Table 11 below.

Table 11: A summary of the Wald test.

Variable	Estimate	Standard Error	Z-Value	P-Value
Intercept	0.9864	0.2147	4.595	4.34×10^{-6}
Classification	0.5898	0.1125	5.241	1.60×10^{-7}

4.3.2 Likelihood Ratio Statistics

The logistic regression null deviance is 33.7861 with degrees of freedom = 3 and the residual deviance is 1.0259 with degrees of freedom = 2. The likelihood ratio statistic for testing H_o: $\beta = 0$ is 33.7861 – 1.0259 = 32.7062 with 3-2=1 degrees of freedom. This also provides strong evidence of a positive effect of classification level on college enrollment.

5. Discussion

The overall objective of this study is to examine the relationship between advanced math courses and college enrollment of our program participants. Results show that classification level has a statistically significant positive effect on college enrollment. Higher estimated probabilities of college enrollment result from more advanced math courses. Results show that a student completing an Advanced III course has a 97% probability of enrolling in college. These results suggest that if students focus on completing at least an Advanced II course they will enroll in college. The work of

Bozick, et. al (2008) shows that students who reach the most advanced math courses such as Pre-Calculus and Calculus before leaving high school are more likely to learn the most advance skills and concepts.

5.1 Relationship Between Math and College Enrollment

Mathematics is a controlling factor for future educational and occupational opportunities. The subject is important in improving the economic and social conditions of students from disadvantage backgrounds. (Byun, et, al. 2015) Studies suggest that inequities in access to and taking advanced math are central differences in college enrollment of students from racial/ethnic minority, low socioeconomic status, and English Language Learner backgrounds. (Byun, et al., 2015). Our program participants fall within these categories and the availability of math courses is different amongst the high schools. Not all program participants have access to advanced math courses, such as Pre-Calculus, available to them. The students can take the subject through dual enrollment, or the Advanced Course Network. However, accessibility for our program participants is challenging.

A report published in 2000 by the U.S. Department of Education, National Center for Education Statistics discusses the rate at which students completed advanced-level high school mathematics courses had a direct bearing on whether or not they enrolled in a 4-year college within two years of graduating from high school. The relationship was especially evident for first-generation students: nearly two-thirds (64 percent) who completed any advanced courses enrolled, compared with about one-third (34 percent) who completed courses through Algebra II. We hope this research will aid in future accessibility for our program participants and help them become the first in their families to go to college.

5.2 Limitations of the Study

This study had several limitations. First being that the College Crusade didn't have access to all the program participant's earned grade information. Though parents of program participants sign a waiver form allowing the College Crusade access to their child's information, some schools do not honor it. Particularly, the charter and private schools. Because of this, smaller sample sizes result. The original cohort had 1,215 students, but we only had grade information for 898 program participants. Another issue regarding the lack of data is the students who completed advanced math courses via dual enrollment or Advanced Course Network were not counted in this analysis.

Second, the courses were not grouped together using a universal code from the State of Rhode Island. Instead, the Gear-Up categories were used to apply the classification levels. The limitation of this method is that any courses with the same or similar titles were lumped together. Each school district may offer the same course Algebra II, but the difficulty level may differ for each. Probability and Statistics in one district may be more advanced than the same titled course in another school. In other words, some of the Advanced 1courses may belong in the Advanced II category pending a curriculum study.

Finally, the sample was not random. Students that enroll in our programs have to have certain requirements: eligible for free/reduced lunch and/or be a first-generation college student. The majority of our program participants are low-income. Also, race/ethnicity was not added to the logistic regression model because the sample is 56% Hispanic or Latino, resulting in bias.

5.3 Future Work

The research findings have supported initial efforts at the College Crusade of Rhode Island to provide access to higher-level math courses for students. The organization will provide dual enrollment opportunities for students to take Pre-calculus and Calculus at the college level and has established a partnership with the Rhode Island Department of Education to create a math readiness program for students. The College Crusade of Rhode Island will also develop programming to prepare students for college placement testing, with a large focus on mathematics. While this is a great start, more work needs to be done.

Future work may include looking the relationship between advanced math courses and the type of college a program participant is enrolled in. Does more advanced math course taking ensure higher estimated probabilities of enrollment in a 4-year institution versus a 2-year? Future work may also include looking at the relationship between advanced math courses and college persistence. Students that enrolled in college, do they persist? If so, what advanced math course did the complete by grade 12? Knowing the answers to these questions would allow for a fine-tuned college access plan for our program participants.

While this study had numerous limitations and there is much room for future work, this preliminary study does indicate that a relationship exists between advanced math course taking and college enrollment. The U.S. Department of Education (2000) showed the advantage of taking advanced mathematics courses was particularly important for first-generation students. Those who completed any advanced mathematics courses in high school enrolled in 4-year colleges at nearly double the rate as those who completed mathematics courses through Algebra II. All relevant stakeholders should take note of math course taking patters of our students and provide guidance to ensure college access and completion.

References

Agresti, Alan. An Introduction to Categorical Data Analysis. Wiley, 2019.

Bilder, Christopher R., and Thomas Loughin. Analysis of Categorical Data with R. CRC Press, 2015

Bozick, R., and Ingels, S.J.. Mathematics Coursetaking and Achievement at the End of High School: Evidence from the Education Longitudinal Study of 2002 (ELS:2002) (NCES 2008-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC., 2008

Byun, Soo-Yong et al. "Advanced Math Course Taking: Effects on Math Achievement and College Enrollment." Journal of experimental education vol. 83,4 (2015): 439-468. doi:10.1080/00220973.2014.919570

Long, M. C., Iatarola, P., & Conger, D. Explaining gaps in readiness for college-level math: The role of high school courses. Education, 4(1), 1–33., 2009.

Long, Mark C.; Conger, Dylan; Iatarola, Patrice. Effects of High School Course-Taking on Secondary and Postsecondary Success, *American Educational Research Journal*, v49 n2 p285-322 Apr, 2012

National Student Clearinghouse, “StudentTracker.” www.studentclearinghouse.org/colleges/studenttracker/. Retrieved March 9, 2020.

U.S. Department of Education, National Center for Education Statistics. Mathematics, Foreign Language, and Science Course-taking and the NELS:88 Transcript Data, NCES 2003–01, by David T. Burkam and Valerie E. Lee. Project Officer: Jeffrey Owings. Washington, DC: 20030

U.S. Department of Education. National Center for Education Statistics. Mapping the Road to College: First-Generation Students’ Math Track, Planning Strategies, and Context of Support, NCES 2000–153, by Laura Horn and Anne-Marie Nuñez. Project Officer: Larry Bobbitt. Washington DC: 2000.