Second-Career Teachers:
How holding a STEM job prior to teaching affects students' performance and entry into STEM fields

Kiersten Barr
Quantitative Methods in the Social Sciences
Columbia University
Master's Thesis
May 2022

Table of Contents

Abstract

Over the past 3 decades, it has become clear that the US ranks much lower than our peer countries in math and science education. As a result, research into what makes an effective STEM teacher has exponentially increased. This paper explores a teacher characteristic not yet investigated by large scale quantitative studies: how math, science, and STEM teachers having math-, science-, and STEM-related jobs prior to teaching affects students' performance in math, science, and STEM and entry into STEM fields in their post-secondary careers. While this study does not find significant positive or negative effects associated with STEM teachers having STEM-related jobs prior to teaching on student performance in STEM or entry into STEM-fields, it confirms findings found elsewhere, such as the effects of student gender, race, and socioeconomic status and teacher experience on student entry into STEM fields

Introduction and Literature Review

Research Question and Rationale for Study

A Nation at Risk, published by the National Commission on Excellence in 1983, was the first report of its kind to bring widespread attention to the status of science and mathematics education in the United States. It issued a warning: the innovation of the US in technology and the sciences were being overshadowed by other foreign competitors (National Commission on Excellence, 1983). The report issued a clear solution: we, as a nation, must commit ourselves to education reform for the goal of excellence, especially in mathematics and the sciences.

Since then, funding for STEM (Science, Technology, Engineering, and Mathematics) education research has increased dramatically and with it came research into all aspects of STEM education (Suter and Camilli, 2018). Studies have been conducted into teaching STEM fields using multiple visual representations, problem-solving methods, body-based tasks, to name a few of the many angles STEM research has taken. These studies, and therefore the funding, will prove worthless if there is not a qualified, effective STEM teacher workforce. Many studies have researched how specific teacher characteristics affect teacher effectiveness and student performance. One topic that is missing from this group of research is the effect of second-career STEM teachers, specifically those who have held STEM-related jobs prior to teaching, on student performance. While research exists on second-career teachers, these articles tend to be case-study based in nature. A large, quantitative analysis of the effectiveness of STEM teachers who held careers in STEM prior to teaching is missing from the current literature.

Literature Review

Do Teacher Characteristics Impact Student Performance?

There are many published studies that investigate how different teacher characteristics, such as subject matter preparation, teaching experience, postsecondary degrees, and certification status, impact student performance. There are some studies that yield conflicting results, while others build upon the ideas of earlier researchers.

Subject Matter Preparation / Content Knowledge. Many studies suggest that STEM students learn more from teachers with more STEM subject-matter preparation. In one such study, using data from the Longitudinal Survey of American Youth (LSAY), Monk (1994) found that the subject matter preparation of teachers, measured by the number of courses taken in the subject being taught, is associated with high school students' mathematics and science test score gains. Students of teachers with more subject matter preparation yielded higher gains in mathematics and science (Monk, 1994).

Rather than using courses taken in mathematics to measure mathematical content knowledge, Rowan, Chiang, and Miller (1997) use a one item math quiz included in the National Educational Longitudinal Study of 1988 (NELS:88) to predict student learning gains. The researchers found that students whose teachers answered the item correctly yielded larger mathematics gains between 8th and 10th grades.

Rockoff, Jacob, Kane, and Staiger (2011) argue that it is not pure mathematical knowledge that is most important, but *mathematical knowledge for teaching*. "Mathematical knowledge for teaching involves the ability to explain difficult mathematical concepts in multiple ways and to describe the intuition behind mathematical reasoning" (Rockoff, et al., 2011). Using a multiple-choice measurement tool of mathematical knowledge for teaching,

developed by researchers at the University of Michigan (Hill, Rowan, and Ball, 2005), Rockoff, et al. found that mathematical knowledge for teaching was strongly related to increases in student achievement. This finding was significant at the .09 level.

Postsecondary Degrees. Using NELS:88 data, Chiang (1996), studied the interaction between teachers majoring in mathematics at the undergraduate or graduate level and students' mathematical achievement in the 10th grade. Chiang (1996) found that teachers holding either bachelor's or master's degrees in mathematics yielded a significant increase in students' 10th-grade mathematical achievement. Interestingly, teachers with both a bachelor's and master's degree in mathematics were not more effective at improving students' mathematics achievement than teachers with only one degree. Additionally, Chiang found that teachers with only master's degrees in mathematics did not yield students with higher levels of mathematical achievement than teachers with purely bachelor's degrees (Chiang, 1996).

Chiang's (1996) findings were confirmed in another analysis of the NELS:88 data; Goldhaber and Brewer (1997) found a positive relationship between teachers' postsecondary degrees and students' mathematical performance. Teachers who held higher education (bachelor's or master's) degrees in mathematics yielded higher-performing students than teachers who were teaching out of field (Goldhaber and Brewer, 1997).

Certification Status

Goldhaber and Brewer confirmed their 1997 findings in their 2000 study when analyzing the relationship between postsecondary degrees, certification status of teachers, and students' performance in mathematics and science. Also, they found that teachers who were certified to teach mathematics but did not hold a degree in mathematics yielded students who did not

perform as well as students with teachers who held a degree in mathematics and were certified (Goldhaber and Brewer, 2000).

Certification Status. Goldhaber and Brewer also took their analysis a step further to investigate how different types of certifications (standard, emergency, probationary, public, etc.) affect student outcomes. Teachers who were not certified to teach mathematics or those who held private school certification yielded lower-performing students than teachers with standard, probationary, or emergency certification. On average, Mathematics teachers with standard certifications, as opposed to private school certifications or certifications out of subject, saw a 1.3-point increase in their students' mathematics scores. Additionally, researchers found that students of mathematics teachers with emergency certification fare equally to those of teachers with standard certification (Goldhaber and Brewer, 2000).

Pedagogical Content Knowledge. As previously discussed, subject coursework has been proven to yield positive effects on student performance. Can we say the same about pedagogical coursework? Monk (1994) found that coursework in pedagogy had positive effects on student learning gains. Additionally, in some years investigated, pedagogy coursework was more influential on student achievement than content area preparation.

In contrast to Monk's findings, Goldhaber and Brewer (2000) found that having a bachelor's degree in education, and therefore more courses in pedagogy, had a statistically significant negative impact on students' mathematics performance. Having a bachelor's degree in education does not affect science outcomes. The authors suggest this is due to the fact that college students majoring in education are drawn from the lower end of the student achievement spectrum.

Non-Traditional and Second Career Teachers

Across the United States, states are seeing STEM teacher shortages. According to data published by the Department of Education, in the 2019-2020 school year, 45 of the 50 U.S states saw mathematics teacher shortages across the entire state or in at least one geographic region (US Department of Education, 2020). In the sciences, 43 states have seen the same in the 2019-2020 school year (US Department of Education, 2020). Recruitment of STEM professionals, i.e. individuals holding jobs in STEM fields, has been considered a way in which the US could ease the STEM teacher shortage (Resta, et al., 2001). As a result, alternative pathways into teaching have increased dramatically (Haselkorn and Hammerness, 2008). For example, in 2012, New York State adopted regulations that allow individuals with postsecondary teaching experience and graduate degrees in their field of study to obtain a teaching license through multiple pathways (NYSED, 2012).

With more pathways to enter the field of education, more STEM professionals are choosing to enter the field of education and teachers are entering the career later in life (Haselkorn and Hammerness, 2008). Researchers have found common themes as to why STEM professions would change careers. In a 1994 qualitative study into second career teachers based on interviews with 50 individuals who change careers to go into teaching, Freidus found that most career-changers are on a mission. They are usually intrinsically motivated to teach, wanting to "pay back" and make the world a better place (Freidus, 1994). There is also evidence to suggest that second-career teachers bring a valuable, different perspective to the field of education. In a qualitative case study of STEM career changers, Grier and Johnston (2009), found that second-career teachers showed very strong "people skills," which they attributed to maturity, life experience, and previous work experience. STEM career changers also felt

confident in their abilities to make real-world connections to the material for their students (Grier and Johnston, 2009).

Overall, quantitative analyses of teacher effectiveness in relation to first-career or second-career teachers are lacking. I was able to find one article which investigated this phenomenon. Boyd et al. (2011) surveyed 4,016 new NYC elementary and middle public school teachers during the 2004-2005 academic year. Researchers asked teachers if they had "worked in a profession other than teaching full-time after completing their undergraduate degree: not at all, less than one year, one to two years, three to five years, or six or more years" (Boyd et al., 2011). Using data provided from NYSED, researchers were able to tie survey responses to student test scores on the ELA and math states tests for grades 4-8. With or without control for teacher characteristics, the researchers found that overall second-career teachers with significant prior experience were less effective than their peers with no experience prior to teaching. The negative effects were largest in mathematics (Boyd, et al., 2011).

Literature Review Conclusion

There is extensive research on the effects of teacher characteristics on student achievement. Researchers have covered subject matter knowledge, pedagogical content knowledge, postsecondary degrees, certification status, and many more topics when investigating teacher effectiveness. One area that is missing from this body of research is how teachers holding STEM jobs prior to teaching affect student achievement in STEM fields. While researchers have investigated second-career teachers, none of these studies focus specifically on secondary STEM teachers, directly tie this to student achievement, and most of these are case studies and qualitative in nature. Boyd et al. (2011) made headway on this topic, but their research was

centered around elementary and middle school teachers and did not limit previous experience to only experience in the subject being taught.

I suspect the lack of research in this field is due to a lack of data surrounding the prior employment of teachers. Fortunately, a newly published longitudinal study (High School Longitudinal Study of 2009 - HSLS:09) conducted by the National Center for Education Statistics included questions about the prior experience of math and science teachers, specifically in STEM fields.

Hypotheses

Given the fact that math teachers who have held math-related jobs prior to teaching are likely to bring more coursework and degrees in their subject, which was proven to have a positive effect on student achievement (Monk, 1994), I have two main hypotheses:

1.) Teachers having prior STEM jobs positively affects student performance in STEM, broken up as follows:

   a. Students with math educators who have held math-related jobs prior to teaching will show higher achievement in high school math.

   b. Students with science educators who have held science-related jobs prior to teaching will show higher achievement in high school science.

   c. Students with STEM educators (combined math and science) who have held STEM-related jobs prior to teaching will show higher achievement in high school STEM.

2.) Teachers having prior STEM jobs positively affects students' entrance into STEM fields, broken up as follows:

    a. Students with math educators who have held math-related jobs prior to teaching will be more likely to enter STEM fields in their post-secondary careers.

    b. Students with science educators who have held science-related jobs prior to teaching will be more likely to enter STEM fields in their post-secondary careers.

    c. Students with STEM educators (combined math and science) who have held STEM-related jobs prior to teaching will be more likely to enter STEM fields in their post-secondary careers.

Methods

Description of Dataset and Variables

Description of Dataset

I am using data from the High School Longitudinal Study of 2009 (HSLS:09) conducted by the National Center for Education Statistics (NCES). This study surveyed about 23,000 randomly selected 9th-graders from 944 U.S. public, private, and Catholic high schools in 2009. Schools had to include 9th and 11th grades to be included in the study. The student sample is representative of the national US 9th grade population in 2009. Students were surveyed throughout their high school and postsecondary years, with data being collected in 2009, 2012, 2013-14, 2016, and 2017-18. Students who dropped out of school were still followed and surveyed at the same time as their peers who stayed in school. In addition to students, their parents, math teachers, science teachers, school administrators, and school counselors were all surveyed. Additionally, the study conducted a self-designed assessment of algebraic concepts to 9th and 11th graders in the study.

In this study, math and science teachers self-reported if they held math- or science-related jobs prior to teaching. This field will make up my independent variable. This dataset is valuable in that it has many variables on which to control: gender, race, socioeconomic status, and school-wide demographics and statistics.

Description of Variables

Key Dependent Variables.

Student Achievement in Math, Science, and STEM. In order to test my first hypothesis, I will be combining the variables listed below into three separate dependent variables in order to test the three versions of my first hypothesis (math, science and STEM).

AchieveMath (Hypothesis 1a):

- X3GPAMATH (renamed HS_GPA_MATH) – Collected from the student's HS transcript. This variable represents the average GPA across all math courses in high school.

- X3THIMATH (renamed Highest_HS_Math) – Collected from the student's HS transcript. This variable represents the highest level of math reached by a student. See Table 1 below for the order / coding of these courses.

Table 1 - Highest_HS_Math (X3THIMATH) Category Labels

Category	Label
0	No Math
1	Basic Math
2	Other Math
3	Pre-algebra
4	Algebra I
5	Geometry
6	Algebra II
7	Trigonometry
8	Other advanced math
9	Probability and statistics
10	Other AP/IB math
11	Precalculus
12	Calculus
13	AP / IB Calculus
-9	Missing
-8	Non-response

- X3TCREDMATH (renamed Credits_MATH) - Collected from the student's HS transcript. This variable represents the total number of credits a student received in HS math courses. Carnegie Units were used. A Carnegie unit is equivalent to a one-year academic course taken one period a day, five days a week.

- X2TXMTH (renamed F1_Math_Score) – A score representing a student's

  mathematical ability using a computer adaptive test given to students in the first follow-

  up of the study (Spring 2012, 11th grade). The test was meant to measure a student's

  algebraic reasoning. Students were given an initial router test, the results of which

  determined the next test (low, moderate, or high ability). This field is the mathematics

  theta score, relative to the achievement of the entire population.

As shown in Table 2 below, these variables have a standardized Cronbach's alpha scale

reliability of 0.80, indicating that *AchieveMath* is a relatively strong indicator. I've chosen to

use the standardized alpha in this case because some of the variables included are on different

scales.

Table 2 – AchieveMath Reliability Statistics

Test scale = mean (standardized items)	
Number of Items in Scale	4
Average Interitem covariance	0.50
Standardized Scale reliability coefficient	0.80

AchieveScience (Hypothesis 1b):

- X3GPASCI (renamed HS_GPA_SCI) – Collected from the student's HS transcript.

  This variable represents the average GPA across all science courses in high school.

- X3THISCI (renamed Highest_HS_Science) - Collected from the student's HS

  transcript. This variable represents the highest level of science reached by a student.

  See Table 3 below for the order / coding of these courses.

Table 3 - Highest_HS_Science *(X3THISCI) Category Labels*

Category	Label
0	No Science
1	General Science
2	Specialty Science
3	Advanced Science
5	AP/IB Science
-9	Missing
-8	Non-response

- X3TCREDSCI (renamed Credits_SCI) - Collected from the student's HS transcript. This variable represents the total number of credits a student received in HS science courses. Carnegie Units were used. A Carnegie unit is equivalent to a one-year academic course taken one period a day, five days a week.

As shown in Table 4 below, these variables have a standardized Cronbach's alpha scale reliability of 0.75, indicating that *AchieveSci* is a relatively strong indicator. I've chosen to use the standardized alpha in this case because some of the variables included are on different scales.

Table 4 – AchieveSci Reliability Statistics

Test scale = mean (standardized items)	
Number of Items in Scale	3
Average Interitem covariance	0.50
Standardized Scale reliability coefficient	0.75

AchieveSTEM (Hypothesis 1c):

- X3GPASTEM (renamed HS_GPA_STEM) – Collected from the student's HS transcript. This variable represents the average GPA across all STEM courses in high school.

- X3THIMATH (renamed Highest_HS_Math) – Explained above. See Table 1 above for the order / coding of these courses.

- X3THISCI (renamed Highest_HS_Science) - Explained above. See Table 3 above for the order / coding of these courses.

- X3TCREDSTEM (renamed Credits_STEM) - Collected from the student's HS transcript. This variable represents the total number of credits a student received in HS STEM courses. Carnegie Units were used. A Carnegie unit is equivalent to a one-year academic course taken one period a day, five days a week. Students with missing or non-response values were removed from the dataset.

- X2TXMTH (renamed F1_Math_Score) – Explained above.

As shown in Table 3 below, these variables have a standardized Cronbach's alpha scale reliability of 0.83, indicating that *AchieveSTEM* is a relatively strong indicator. I've chosen to use the standardized alpha in this case because some of the variables included are on different scales.

Table 5 – AchieveSTEM Reliability Statistics

Test scale = mean (standardized items)	
Number of Items in Scale	5
Average Interitem covariance	0.50
Scale reliability coefficient	0.83

Entrance into STEM. In order to test my second hypothesis, I will be making a variable, EnterSTEM which will indicate if the student entered the STEM profession. As variables indicating entrance into a STEM profession are not broken up by math and science, this variable will be used as the dependent variable for the models testing hypotheses 2a, 2b, and 2c. This variable will be a binary indicatory which will be the combination of the following fields:

- X4OCC1STEM1 (renamed STEM_first_job) – Collected during the second follow-up in 2016, this variable indicates if a student's 1st job out of HS was in a STEM field.

- X4OCCFBMRST1(renamed STEM_occupation) – Collected during the second follow-up in 2016, this variable indicates if a student's occupation in February 2016 or their most recent occupation was in a STEM field.

- X5STEMCRED (renamed STEM_major) – Collected from the student's post-secondary transcripts, this variable indicates if a student received any post-secondary degree in a STEM field.

- X4OCC30STEM1 (Renamed STEM_job_30) – Collected during the second follow-up in 2016, this variable indicates if a student believes their occupation when they are 30 will be in a STEM field.

Key Independent Variables. In order to test all of my hypotheses, my main independent variables will be MathJob (hypotheses 1a and 2a), SciJob (hypotheses 1b and 2b), or STEMjob (hypotheses 1c and 2c). These variables indicate if either the student's 9th-grade math, science, or combined STEM teacher held a math-, science-, or STEM related job prior to teaching. Unfortunately, this data point was only collected in the base year of the study. These fields are explained below:

- M1MATHJOB (renamed MathJob)– Collected during the base year, this variable indicates if a student's 9th-grade math teacher held a math-related job prior to teaching.

- N1SCIJOB (renamed MathJob) – Collected during the base year of the study, this variable indicates if a student's 9th-grade science teacher held a science-related job prior to teaching.

- STEMjob – a created variable combining the two variables above. This field indicates if a student had a 9th-grade math teacher who held a math-related job prior to teaching or if they had a 9th-grade science teacher who held a sceicne-realted job prior to teaching.

Key Control Variables. Below are the key control variables which I will include in all my models:

- X1SEX (renamed student_male) – a student's sex. Male and female students enter STEM fields and succeed in STEM fields at different rates. This will serve as an important control variable as genders tend to ender STEM at different rates.

- X1RACE (renamed student_race) – a student's race. Similarly to sex, students of different races enter STEM fields at different rates. Unknown values were categorized as unknown

Table 6 – student_race Category Labels

Category	Label
1	White
2	Asian
3	Black/African-American, non-Hispanic
4	Hispanic, no race specified or race specified
5	More than one race, non-Hispanic
6	Native Hawaiian/ Pacific Islander, non-Hispanic
7	American Indian / Alaska Native, non-Hispanic
8	Unknown

- X1SES (renamed student_ses) – a variable representing a student's socioeconomic status. This is a composite variable constructed by the survey to measure socioeconomic status. This variable was calculated using parent/guardians' education, occupation, and family income.

- M1ALTCERT and N1ALTCERT (renamed MATH_altcert and SCI_altcert) – these fields, collected during the base year of the study, indicate if 9th-grade math or science teacher entered the teaching program through an alternative certification program. These variables will be included to ensure that any observed effects are not due to teacher certification status.

- M1MATHYRS912 and N1SCIYRS912 (renamed MATH_years and SCI_years) – these fields, collected during the base year of the study, indicate the number of years the 9th-grade math and science teacher has taught 9-12 math and science. Similar to alternative certifications, years teaching could impact students' performance in STEM.

Limitations to Data

While this dataset contains the very important datapoint of whether or not a math or science teacher held a math- or science-related job prior to teaching, this datapoint is only available for a student's 9th-grade math and science teacher. This will significantly narrow any

conclusions I can make from results and it will likely prevent me from seeing any significant results. I will still continue with this research nonetheless. Additionally, I foresee issues with the lack of teachers that fall into this category of STEM career changes. I found that only 16.5% of the students in the study had a 9th-grade STEM teacher who held a STEM-related job prior to teaching. Additionally, this data point was self-reported by the teacher, and did not ha e

Additionally, while the study tried to follow all recipients over time, there is a lot of missing data points after the first round of data collection. I explain below how I will deal with this.

Dealing with Missing Data

Although the dataset is massive at first glance, there are a lot of missing values and non-respondents, especially in the later rounds of the longitudinal study. In order to move forward with this analysis, I had to decide how to deal with this missing data. I chose to impute the mean for some fields. Others, I removed missing records entirely.

In order to include as many records as possible in all of my models, I created three different datasets, each with missing records removed for different variables depending on which were included in the regressions. Below are the fields in which I chose to remove missing data entirely to create the three data sets. These were fields where imputation was impossible. Prior to removing these fields, I had 23,503 records in my dataset.

Table 7: Fields in which NA Values were Removed

Dataset A Hypotheses 1a and 2a	Dataset B Hypotheses 1b and 2b	Dataset C Hypotheses 1c and 2c
Math_Job Math_altcert Enter_STEM	Sci_Job Sci_altcert Enter_STEM	STEM_Job Math_altcert Sci_altcert Enter_STEM
Final N: 12,926	Final N: 11,967	Final N: 9,960

Table 8 below contains the fields in which I chose to impute the mean and their associated statistics. Overall, computing to the mean did not drastically change the mean and median, and therefore the overall distribution, for these fields.

Table 8: Field Summaries Prior and Post Imputation

Field	Prior to Imputation				Post Imputation			
	Non N/A Count	Mean	SD	Median	N	Mean	SD	Median
Dataset A: Hypotheses 1a and 2a								
HS_Math_GPA	12,363	2.479	0.969	2.500	12,926	2.474	0.948	2.500
Highest_HS_Math	12,366	8.575	3.087	8	12,926	8.557	3.021	8.000
HS_Math_Cred	12,385	3.718	1.169	4.000	12,926	3.712	1.145	4.000
Math_Years_Taught	12,902	10.316	8.552	8	12,926	10.316	8.545	8.000
F1_Math_Score	12,010	0.834	1.138	0.784	12,926	0.826	1.098	0.717
Student_SES	12,200	0.128	0.787	0.071	12,926	0.124	0.765	0.054
Dataset B: Hypotheses 1b and 2b								
HS_Sci_GPA	11,493	2.581	0.965	2.500	11,967	2.574	0.946	2.500
Highest_HS_Sci	11,501	2.179	1.472	2	11,967	2.172	1.444	2.000
HS_Sci_Cred	11,518	3.477	1.238	3.500	11,967	3.468	1.216	3.500
Sci_Years_Taught	11,963	10.472	7.904	8	11,967	10.472	7.903	8.000
student_ses	11,311	0.155	0.788	0.098	11,967	0.150	0.766	0.054
Dataset C: Hypotheses 1c and 2c								
HS_Math_GPA	9,589	2.509	0.972	2.500	9,960	2.503	0.954	2.500
HS_Sci_GPA	9,589	2.583	0.968	2.500	9,960	2.577	0.951	2.500
Highest_HS_Math	9,592	8.685	3.092	8	9,960	8.665	3.036	8.145
Highest_HS_Sci	9,592	2.176	1.469	2	9,960	2.170	1.441	2.000
Tot_HS_STEM_Cred	9,608	7.870	2.428	8.000	9,960	7.856	2.386	8.000
Math_Years_Taught	9,930	10.484	8.615	8	9,960	10.483	8.602	8.000
Sci_Years_Taught	9,948	10.552	7.911	8	9,960	10.552	7.906	8.000
F1_Math_Score	9,309	0.878	1.132	0.832	9,960	0.867	1.095	0.748
student_ses	9,430	0.160	0.790	0.104	9,960	0.154	0.769	0.054

Imputing these fields and removing missing data did affect the overall distribution of some data points. For example, the percentage of students with 9th grade STEM teachers who held STEM related jobs prior to teaching went from 31.9% to 45.1%. See figure 1 and figure 2 for the full comparison of distribution for this field. This is a huge jump and likely adds some bias to the STEM models.
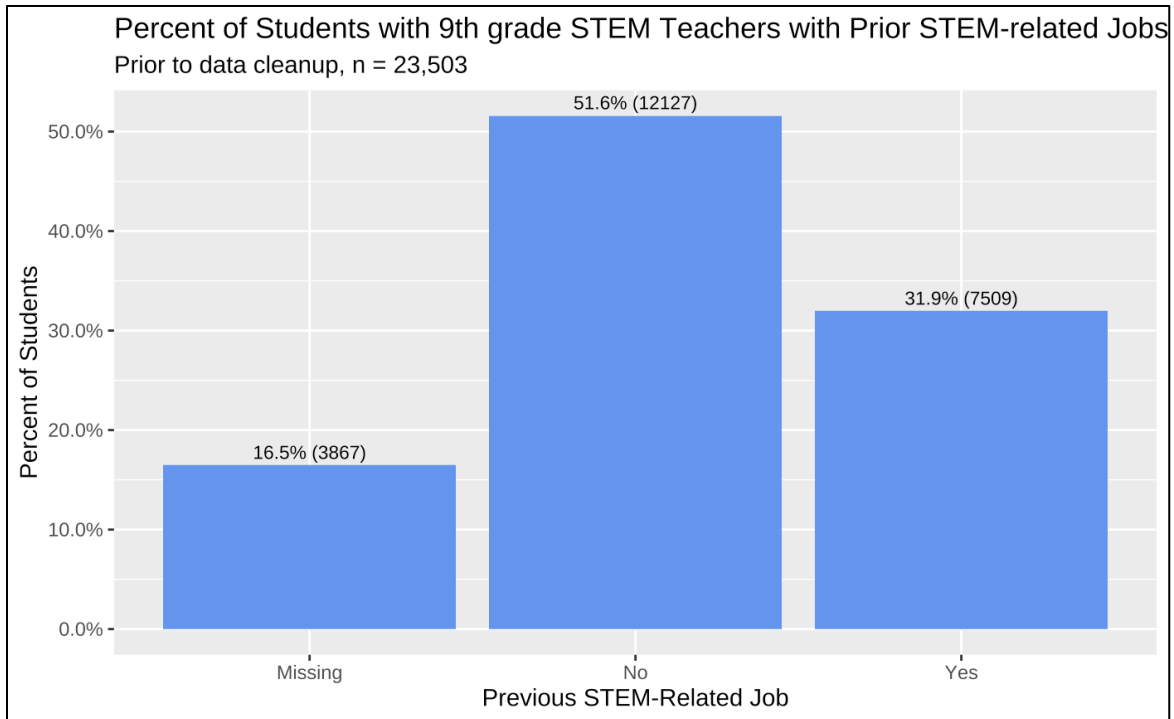

Figure 1: Percent of Students with 9th-grade STEM Teachers with Prior STEM-related Jobs, Pre Data Cleanup
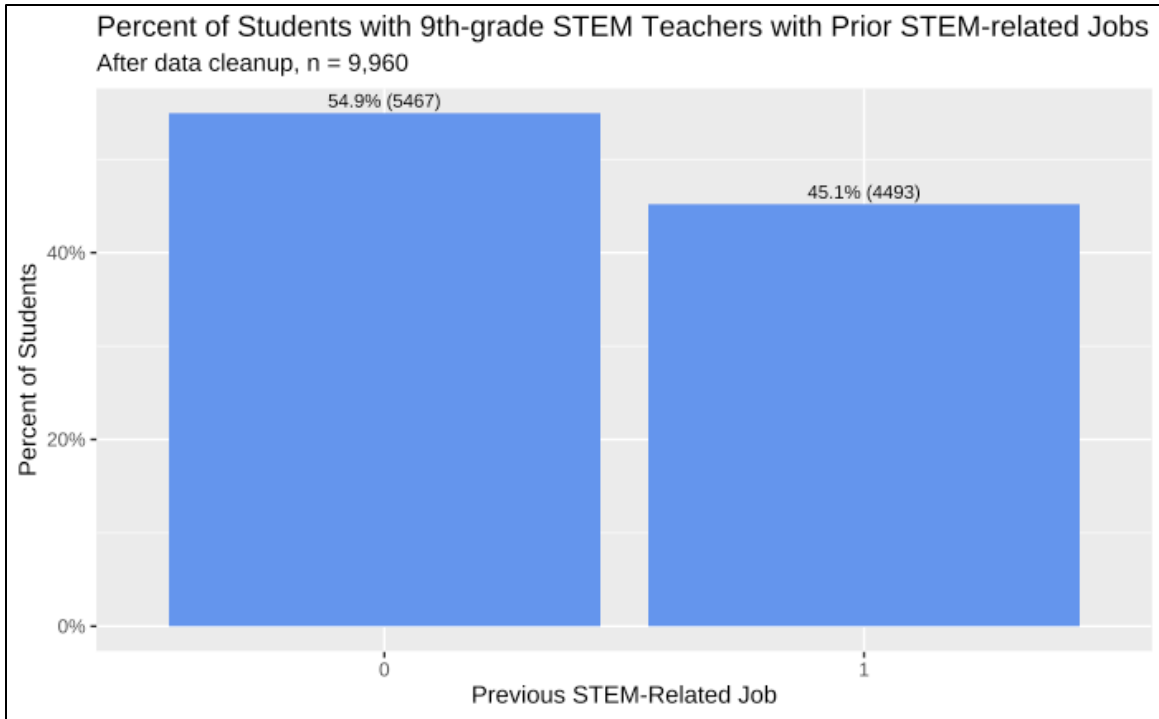
Figure 2: Percent of Students with 9ᵗʰ-grade STEM Teachers with Prior STEM-related Jobs, Post Data Cleanup

Additionally, after removing missing data and imputing fields, the percentages of students who entering STEM fields in their post-secondary careers changed from 6% to 8.4%. This is a slight change and is unlikely to affect the results of the models. See figure 3 and figure 4 below.
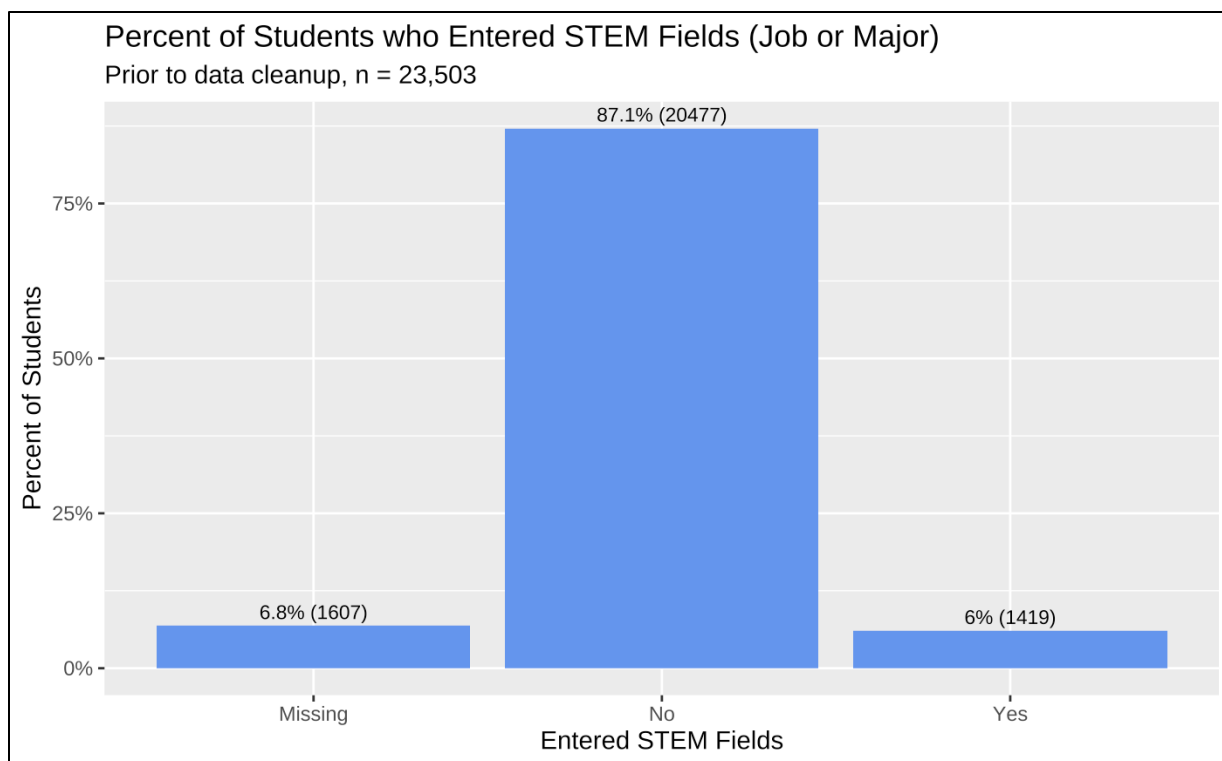
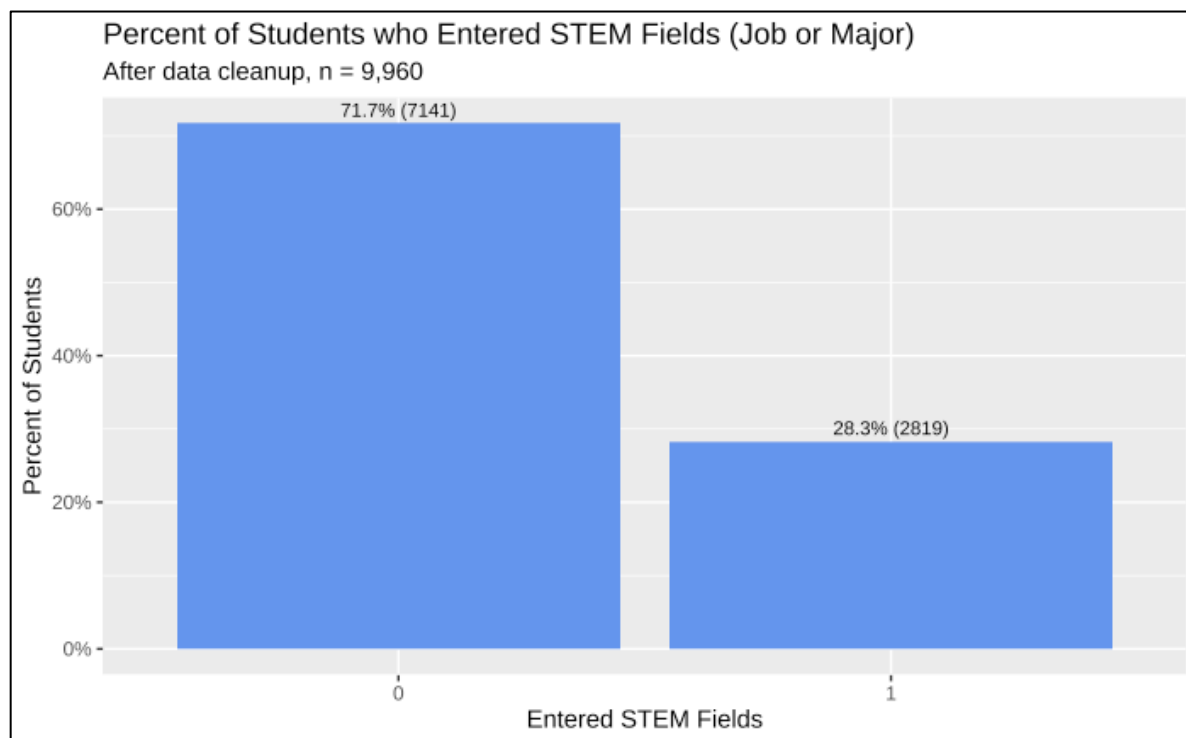*Figure 3: Percent of Students who Entered STEM fields, pre data cleanup*

Figure 4: Percent of Students who entered STEM fields, post data cleanup

Descriptive Statistics

Below are the descriptive statistics for my variables after removing missing datapoints across the entire data set.

Table 9: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
Dataset A: Hypotheses 1a and 2a						
MATH_job	12,926	0.182	0.386	0	0	1
AchieveMATH	12,926	0.000	0.788	-2.978	0.012	2.379
ENTERstem	12,926	0.276	0.447	0	0	1
student_male	12,926	0.480	0.500	0	0	1
student_ses	12,926	0.124	0.765	-1.930	0.054	2.567
MATH_altcert	12,926	0.187	0.390	0	0	1
Math_Years_Taught	12,926	10.316	8.545	1.000	8.000	31.000
Dataset B: Hypotheses 1b and 2b						
SCI_job	11,967	0.333	0.471	0	0	1
AchieveSCI	11,967	-0.000	0.815	-2.711	-0.096	2.398
ENTERstem	11,967	0.282	0.450	0	0	1
student_male	11,967	0.481	0.500	0	0	1
student_ses	11,967	0.150	0.766	-1.930	0.054	2.567
SCI_altcert	11,967	0.284	0.451	0	0	1
Sci_Years_Taught	11,967	10.472	7.903	1.000	8.000	26.000
Dataset C: Hypotheses 1c and 2c						
STEM_job	9,960	0.451	0.498	0	0	1
AchieveSTEM	9,960	0.000	0.771	-2.757	-0.038	2.262
ENTERstem	9,960	0.283	0.450	0	0	1
student_male	9,960	0.480	0.500	0	0	1
student_ses	9,960	0.154	0.769	-1.930	0.054	2.567
MATH_altcert	9,960	0.175	0.380	0	0	1
SCI_altcert	9,960	0.284	0.451	0	0	1
Math_Years_Taught	9,960	10.483	8.602	1.000	8.000	31.000
Sci_Years_Taught	9,960	10.552	7.906	1.000	8.000	26.000

Prior to combining M1MATHJOB and N1SCIJOB into STEMJOB for dataset C, only 18% (dataset A) and 33% (dataset C) of students had 9th-grade teachers who had math- or science- related jobs prior to teaching, respectively. After combining these variables into STEMJOB for dataset C, nearly 45% of students had 9th-grade STEM teachers who held STEM-related jobs prior to teaching.

In regards to my key dependent variables, we can see that about 28% of students included are considered to have entered stem (EnterSTEM) in all three datasets. The Achievement in Math, Science, and STEM variables (AchieveMath, AchieveSCI, and AchieveSTEM) are centered around zero. These variables are composed of HS_STEM_GPA, Highest_HS_Math, Highest_HS_Sci, and Tot_HS_STEM_Cred.

In regards to my control variables, 50% of the population is male in all three datasets. The average SES score varies slightly across the three datasets, it varies from 0.12 (dataset A) to 0.15 (dataset C). For the variables corresponding to the students' 9th-grade math and science teachers, 18% of students had a 9th-grade math teacher who entered teaching through an alternative certification program in dataset C, and 28% of students had a 9th-grade math teacher who did the same. The average number of years taught for both math and science teachers was between 10 and 11 years in all three regressions. Both the math and science years taught variables have fairly high standard deviations of 8.67 and 7.9, respectively.

Research Design

Hypothesis 1a

In order to test my first hypothesis, I will be running an ordinary least squares regression. AchieveMath, explained above, will be my dependent variable; Math_job will be my key independent variable to pay close attention to. I will be including gender, race, ses, math alternative certifications, and number of years math teachers taught as control variables. This model will take the form of below.

Model 1a: AchieveMath = β0 + β1(Student_Male) + β2(Student_Race) + β3(Student_SES) +
β4 (Math_altcert) + β5 (Math_Years_Taught) + β6 (Math_job)

Hypothesis 1b

In order to test my second hypothesis, I will be running an ordinary least squares regression. AchieveSci, explained above, will be my dependent variable; Sci_job will be my key independent variable to pay close attention to. I will be including gender, race, ses, science alternative certifications, and number of years science teachers taught as control variables. This model will take the form of below.

Model 1b: AchieveSci = β0 + β1(Student_Male) + β2(Student_Race) + β3(Student_SES) +
+ β5 (Math_Years_Taught) + β6 (Sci_job)

Hypothesis 1c

In order to test my third hypothesis, I will be running an ordinary least squares regression. AchieveSTEM, explained above, will be my dependent variable; STEMjob will be my key independent variable to pay close attention to. I will be including gender, race, ses, math

and science alternative certifications, and number of years taught for math and science teachers as control variables. This model will take the form of below.

Model 1c: AchieveSTEM = β0 + β1(Student_Male) + β2(Student_Race) + β3(Student_SES) + β4(Math_altcert) + β5(SCI_altcert) + β6 (Math_Years_Taught) +β7 (Sci_Years_Taught) + β8 (STEMjob)

Hypothesis 2a

In order to test my fourth hypothesis, I will be running binary logit regression for ease of interpretability. EnterSTEM will be my binary independent variable; Math_job will be my key independent variable to pay close attention to. I will be including gender, race, ses, math alternative certifications, and number of years math teachers taught as control variables. This model will take the form of below:

Model 2a: EnterSTEM= β0 + β1(Student_Male) + β2(Student_Race) + β3(Student_SES) + β4(Math_altcert) + β5 (Math_Years_Taught) + β6 (Math_job)

Hypothesis 2b

In order to test my fifth hypothesis, I will be running binary logit regression for ease of interpretability. EnterSTEM will be my binary independent variable; Sci_job will be my key independent variable to pay close attention to. I will be including gender, race, ses, science alternative certifications, and number of years science teachers taught as control variables. This model will take the form of below.

Model 2b: EnterSTEM= β0 + β1(Student_Male) + β2(Student_Race) + β3(Student_SES) + β5(SCI_altcert) +β6 (Sci_Years_Taught) + β7 (Sci_job)

Hypothesis 2c

In order to test my sixth hypothesis, I will be running binary logit regression for ease of interpretability. EnterSTEM will be my binary independent variable; STEMjob will be my key independent variable to pay close attention to. I will be including gender, race, ses, math and science alternative certifications, and number of years taught for math and science teachers as control variables. This model will take the form of below:

Model 2c: EnterSTEM= β0 + β1(Student_Male) + β2(Student_Race) + β3(Student_SES) +
   β4 (Math_altcert) + β5(SCI_altcert) + β6 (Math_Years_Taught) +β7
   (Sci_Years_Taught) + β8 (STEMjob)

Results

Hypotheses 1a, 1b, and 1c

In order to initially test my first hypothesis, that students with STEM teachers who held STEM related jobs prior to teaching showing higher achievement in STEM, I ran the linear model shown in table 10.

Table 10: Models 1a, 1b, and 1c – Hypothesis 1 Results

	Dependent variable:		
	AchieveMATH	AchieveSCI	AchieveSTEM
	(1a)	(1b)	(1c)
Math_job	0.026		
	(0.017)		
Sci_job		-0.017	
		(0.015)	
STEM_job			0.005
			(0.014)
student_male	-0.090***	-0.142***	-0.055***
	(0.012)	(0.013)	(0.014)
student_race.factAsian	0.423***	0.529***	0.517***
	(0.024)	(0.025)	(0.026)
student_race.factBlack	-0.243***	-0.229***	-0.223***
	(0.022)	(0.024)	(0.025)
student_race.factHispanic	-0.097***	-0.089***	-0.082***
	(0.019)	(0.021)	(0.021)
student_race.fact2+ races	-0.083***	-0.026	-0.016
	(0.023)	(0.025)	(0.025)
student_race.factNative_Haw/PI	-0.149	-0.136	-0.125
	(0.092)	(0.108)	(0.106)
student_race.factAmer_Ind	-0.292***	-0.304***	-0.307***
	(0.077)	(0.082)	(0.081)
student_race.factUnknown	-0.081**	-0.022	-0.049
	(0.036)	(0.040)	(0.040)
imputed_student_ses	0.381***	0.361***	0.385***
	(0.008)	(0.009)	(0.009)
MATH_altcert	-0.017		-0.021
	(0.017)		(0.018)

imputed_Math_Years_Taught	0.006***		0.006***
	(0.001)		(0.001)
SCI_altcert		0.022	0.005
		(0.016)	(0.016)
imputed_Sci_Years_Taught		0.003***	0.004***
		(0.001)	(0.001)
Constant	-0.049***	-0.026	-0.140***
	(0.014)	(0.016)	(0.019)
Observations	12,926	11,967	9,960
R^2	0.211	0.194	0.239
Adjusted R^2	0.211	0.193	0.238
Residual Std. Error	0.701 (df = 12913)	0.732 (df = 11954)	0.673 (df = 9945)
F Statistic	288.238*** (df = 12; 12913)	239.056*** (df = 12; 11954)	222.638*** (df = 14; 9945)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

As explained previously, AchieveSTEM is a composite variable constructed from HS STEM GPA, total HS STEM credits, highest math course taken, and highest science course taken. This composite variable has an alpha of 0.83. While table 10 shows that there is a positive relationship between a math teacher having a math-related job prior to teaching and student performance in math, this relationship is not a significant one. The same is true for the relationship between a STEM teacher having a STEM-related job prior to teaching and student performance in STEM; the relationship is positive, but not significant. Interestingly, regression 1b shows that there is a negative relationship between a science teacher having a science-related job prior to teaching and student performance in science, but again this relationship is not significant. These three results disprove hypotheses 1a, 1b, and 1c,

My regressions show that achievement in STEM is significantly affected by the student's gender, race, socioeconomic status, and the years that the student's 9th grade math and science

teachers have taught math and science. Net of all other variables, being male decreases your achievement in STEM by -0.06, on average. This is statistically significant at the p< 0.001 level. This relationship is evident in the math and science regression as well. Additionally, having a 1-point higher socioeconomic status leads to a 0.38 increase in achievement in STEM, on average, net of all other variables. This is statistically significant at the $p < 0.001$ level. This relationship is similarly shown in the math and science regressions as well. Also, having a math or science teacher with more years of experience statistically impacts your performance in STEM. For each year more the student's 9th grade math teacher has taught 9-12 math, achievement in STEM increases 0.005, on average, net of all other variables. For science teachers, the same change results in a 0.004 increase in STEM achievement, on average, net of all other variables.

In regards to race, white students are the reference category for the race variables. Asian students are the only racial group to perform significantly better than white students in STEM. On average, Asian students perform 0.52 higher than white students in STEM, net of all other variables. This finding is statistically significant at the $p<.001$ level. All of the other races perform worse than white students in STEM, some of which are statistically significant. On average, net of all other variables, Black students perform 0.23 ($p <0.001$) points worse, Hispanics (other races) perform 0.08 ($p<0.001$) worse, and American Indians/Alaskan Natives perform 0.31 ($p<0.001$) points worse than white students. Similar results are evident in the math and science regressions as well.

This model has an adjusted R^2 of .238, meaning that we are explaining 23.8% of the variation in STEM achievement with this model.

Hypotheses 2a, 2b, and 2c

In order to test my second hypothesis, that students with math, science and STEM teachers who held math-, science-, and STEM-related jobs prior to teaching will enter STEM fields in their post-secondary careers more often, I ran the logit models shown in table 11. The odds ratios associated with each variable are shown in table 12.

Table 11: Models 2a, 2b, and 2c – Hypothesis 2 Results

	Dependent variable:		
	ENTERstem		
	(2a)	(2b)	(2c)
MATH_job	0.032		
	(0.054)		
SCI_job		0.033	
		(0.046)	
STEM_job			0.045
			(0.047)
student_male	-0.472***	-0.465***	-0.466***
	(0.040)	(0.042)	(0.046)
student_race.factAsian	0.554***	0.676***	0.672***
	(0.070)	(0.071)	(0.079)
student_race.factBlack	-0.057	0.018	-0.011
	(0.073)	(0.075)	(0.084)
student_race.factHispanic	-0.043	-0.025	-0.024
	(0.062)	(0.066)	(0.072)
student_race.fact2+ races	-0.069	-0.013	0.008
	(0.076)	(0.078)	(0.085)
student_race.factNative_Haw/PI	-0.012	-0.008	-0.122
	(0.304)	(0.341)	(0.370)
student_race.factAmer_Ind	0.039	0.030	0.078
	(0.257)	(0.263)	(0.279)
student_race.factUnknown	-0.219*	-0.251*	-0.339**
	(0.125)	(0.132)	(0.149)
imputed_student_ses	0.315***	0.344***	0.322***

	(0.027)	(0.028)	(0.031)
MATH_altcert	0.012		-0.009
	(0.054)		(0.062)
imputed_Math_Years_Taught	0.005**		0.006**
	(0.002)		(0.003)
SCI_altcert		-0.039	-0.074
		(0.049)	(0.053)
imputed_Sci_Years_Taught		0.007**	0.006**
		(0.003)	(0.003)
Constant	-0.892***	-0.913***	-0.959***
	(0.044)	(0.050)	(0.064)
Observations	12,926	11,967	9,960
Log Likelihood	-7,414.832	-6,908.278	-5,758.565
Akaike Inf. Crit.	14,855.660	13,842.560	11,547.130

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 10 – Logit Odds Ratios

	Dependent variable:		
	ENTERstem		
	(2a)	(2b)	(2c)
MATH_job	1.033		
SCI_job		1.034	
STEM_job			1.046
student_male	0.624	0.628	0.627
student_race.factAsian	1.740	1.966	1.958
student_race.factBlack	0.944	1.018	0.989
student_race.factHispanic	0.958	0.975	0.976
student_race.fact2+ races	0.933	0.988	1.008
student_race.factNative_Haw/PI	0.988	0.992	0.885
student_race.factAmer_Ind	1.040	1.030	1.081
student_race.factUnknown	0.804	0.778	0.713
imputed_student_ses	1.370	1.411	1.379
MATH_altcert	1.012		0.991
imputed_Math_Years_Taught	1.005		1.006
SCI_altcert		0.962	0.929

imputed_Sci_Years_Taught		1.007	1.006
Constant	0.410	0.401	0.383

The EnterSTEM variables combine three difference variables into one binary indicator. Any student who had a 1st job out of HS in a STEM field, held a STEM occupation in 2016, received a STEM degree at any point, or thought their job at age 30 would be in a STEM field were considered to have 'Entered STEM' for this analysis. Again, the model disproves hypotheses 2a, 2b, and 2c. Although the direction of the coefficients on Math_Job, Sci_Job, and STEM_Job are in line with my hypotheses, these coefficients are not significant, and therefore having a math-, science-, or STEM- related job prior to teaching does not significantly affect students' entry into STEM fields. However, there are other statistically significant findings in this model on average, net of all other variables, your odds of entering a STEM field are 37% ($p<0.001$) lower for males than for females. Additionally, a student's odds of entering a STEM field go up by 74-95% ($p<0.001$) when the student is Asian, as compared to white students, on average, net of all other variables. I also find that SES significantly affects students' entry into STEM fields. For each point higher the student's SES is, their odds of entering a STEM field go up by 37 - 41% ($p<0.001$). Lastly, similar to the regressions for achievement in STEM, I find that math and science teaching experience significantly affects student's entry into STEM fields. This affect is small but significant. For each more year of experience in teaching math, a student's odds of entering a STEM field go up by 0.5% ($p <0.001$). Likewise, for each more year of experience in teaching science, a student's odds of entering a STEM field go up by 0.6% ($p<0.001$).

Discussion and Conclusion

It was found that alternative certifications in math and science did not have a positive or negative significant effect on student performance in math, science, and STEM and entry into STEM fields. This conflicts with Goldhaber and Brewer's 2000 finding that standard mathematic teaching certification led to a 1.3-point increase in their students' mathematics scores (Goldhaber and Brewer, 2000). This may be a factor of how I define mathematical achievement as a scaled score over multiple variables, while Goldhaber and Brewer only used test scores. For this reason, public and private schools should continue to look for teachers from any pathway. Teachers are still needed now more than ever.

Additionally, this study brings to the surface some key differences in achievement and entry into STEM fields between races and sexes. Why these differences exist and how to mitigate them was outside the scope of this paper, but more research should be performed on these subjects. In addition, it'd be interesting if the sex of a math or science teacher affects student's entry performance in these subject and entry into STEM fields. There also could be an interaction between a students' sex and a teachers' sex, but again, this was outside the scope of this study.

Next, we must remember the limitations to our data. The prior STEM-related job field was only available for a student's 9th-grade teacher math and science teacher. More research is needed to determine how teachers in other grades might affect a student's performance and entry into STEM fields. Additionally, I'd be interested to see how having a STEM teacher who had a STEM-related job prior to teaching for all 4 years of high school affects a student's entry into STEM. Additionally, this field was only available as a binary indicator; no additional

information was available. With more detailed information, we may have been able to see an effect based on certain job types.

  While my hypotheses were proven false, there is still valuable information to come out of this analysis. First, prior math-, science-, and STEM-related jobs may not have an impact on student performance in math, science, and STEM and entry into STEM fields, as predicted, but it is good to know that there is not a significant *negative* effect on performance. Second-career STEM teachers should be comforted by the fact that they will not have a negative effect on student's compared to teachers who did not hold STEM jobs prior to teaching. Additionally, local and state governments can still try to close the STEM teacher shortage by recruiting teachers from standard STEM fields.

References

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., O'Brien, R., & Wyckoff, J. (2011). The effectiveness and retention of teachers with prior career experience. Economics of Education Review, 30(6), 1229-1241.

Capraro, M. M., Capraro, R. M., & Wiggins, B. B. (2000). An Investigation of the Effects of Gender, Socioeconomic Status, Race and Grades on Standardized Test Scores.

Chiang, F. (1996). *Ability, motivation, and performance: A quantitative study of teacher effects on student mathematics achievement using NELS:88 data* (Order No. 9635476). Available from ProQuest Dissertations & Theses Global. (304249881). Retrieved from http://ezproxy.cul.columbia.edu/login?url=https://www-proquest-com.ezproxy.cul.columbia.edu/docview/304249881?accountid=10226

Freidus, H. (1994). Supervision of Second Career Teachers What's Our Line? Distributed by ERIC Clearinghouse.

Grier, J., & Johnston, C. (2009). An Inquiry Into the Development of Teacher Identities in STEM Career Changers. *Journal of Science Teacher Education, 20*(1), 57-75. Retrieved November 10, 2020, from http://www.jstor.org/stable/43156495

Goldhaber, D., & Brewer, D. (1997). Why Don't Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity. *The Journal of Human Resources, 32*(3), 505-523. doi:10.2307/146181

Goldhaber, D., & Brewer, D. (2000). Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129-145. Retrieved November 11, 2020, from http://www.jstor.org/stable/1164392

Haselkorn, D., & Hammerness, K. (2008). Encore Performances: Tapping the Potential of Midcareer

    and Second-Career Teachers. Woodrow Wilson National Fellowship Foundation (NJ1).

Hill, H., Rowan, B., & Ball, D. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on

    Student Achievement. *American Educational Research Journal, 42*(2), 371-406. Retrieved

    November 10, 2020, from http://www.jstor.org/stable/3699380

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers

    and student achievement. *Economics of Education Review*, *13*(2), 125–145.

    https://doi.org/10.1016/0272-7757(94)90003-5

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for

    educational reform. The Elementary School Journal, 84(2), 113-130.

New York State Education Department. (2012). Amendment of Sections 80-1.1, 80-3.3 and 80-

    3.7 and addition of section 80-5.22 of the Commissioner's regulations to allow

    individuals with advanced degrees in science, technology, engineering and mathematics

    (STEM) and related teaching experience at the postsecondary level to obtain a teaching

    certificate in earth science, biology, chemistry, physics, mathematics or in a closely

    related field. Retrieved November 6, 2020, from

    https://www.regents.nysed.gov/common/regents/files/documents/meetings/2012Meetings

    /January2012/112brca2.pdf

Resta, V., Huling, L., & Rainwater, N. (2001). Preparing second-career teachers. Educational

    Leadership, 58(8), 60–63.

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can You Recognize an Effective Teacher

    When You Recruit One? *Education Finance and Policy, 6*(1), 43-74. Retrieved November 10,

    2020, from http://www.jstor.org/stable/educfinapoli.6.1.43