What variables can predict student achievement in science?

Erin Anderson, Richard Paquin Morel, Abigail Stein

Northwestern University
Multidisciplinary Program in Education Sciences

DRAFT
PLEASE DO NOT CIRCULATE WITHOUT AUTHORS' PERMISSION

Executive Summary

Fostering young adults who are prepared and willing to pursue careers in science-related fields is critical for the nation and can improve opportunities for the students, yet female and Black and Hispanic students are often underrepresented in the areas. In this brief, we defined science achievement in three outcomes: participation in science coursework, performance in science classes and the science portion of the ACT, and planning, or the intent to major in a science-related field. Students at North Shore High School show similar patterns to the nation at-large, with boys outpacing girls and White students outpacing Black and Hispanic students in performance and participation, though not in planning to major in science. These gaps exist even after statistical adjustments match students who are similar in their previous performance and their non-science GPA, suggesting that achievement gaps are growing over the course of a student's time at NSHS. Some of this may be due to initial placement in honors biology, as our quasi-experimental model shows that students who are otherwise identical in their previous performance are more likely to take at least one science AP class if they take Honors Biology. To address these findings, NSHS may consider expanding earned honors and other ways of increasing access to these tracks in the initial science class. Initially, NSHS may improve efforts to reduce differences between arriving students at all levels.

Introduction

Science education is fundamental for today's society, helping to promote technological innovation and economic growth (Xie, Fang, & Shauman, 2015). A science literate population supports competitiveness in an increasingly global economy (Committee on Science, Engineering, and Public Policy, 2007). In addition to promoting development and prosperity in a global context, employment in science-related fields brings along high wages and high prestige (Xie, 2015). Thus, increasing science achievement and participation in science fields serve both national economic interests and individual career interests.

Researchers and policymakers have recently sounded the alarm that the US is losing its edge in scientific achievement and competitiveness, with possible economic consequences. Specifically, there is a growing concern over the number of individuals qualified for employment in science-related fields (Committee on Science, Engineering, and Public Policy, 2007). Thus, researchers have increasingly turned their attention to science achievement in the K-12 system (Xie, 2015). Improving achievement in science related field, however, is a necessary, but not sufficient condition for remedying such issues. As Xie and colleagues (2015) note, science, and, more generally, STEM (Science, Technology, Engineering, and Mathematics) education is distinct because it is required for employment in a science field. This requires not only robust achievement in science, but also aspirations for a career in the field, since science careers often require extensive training beyond the high school and undergraduate level.

These problems are exacerbated in part by the composition of the human resources talent pool in science, which remains primarily white and male (Xie et al., 2015). Some research attention has been paid to the underrepresentation of women and minorities in science-related majors and careers (P.A. Muller, Stage, & Kinzie, 2001; Riegle-Crumb, Moore, & Ramos-Wada, 2010; Wang, 2013; Xie et al., 2015). This includes descriptive research documenting differences in achievement and aspirations between genders and ethnicities and inferential research exploring factors that account for those observed differences (Xie et al., 2015). The disparities in achievement and participation in science have two-fold consequences. From a human resources point of view, the underrepresentation of women and minorities contributes to concerns about the available talent pool for science employment. From an equity point of view, this underrepresentation locks out women and minorities from high wage, high-status occupations.

In this report, we provide NSHS with some insights into the following question: What variables can predict student achievement in science? This question can be answered as pertaining to three different aspects of "achievement." In this report we call them the 3 Ps - Performance, Participation, and Planning. Performance includes outcomes such as standardized test scores and GPA and is perhaps the most common way we think about achievement. However, achievement could also be conceptualized as participation, such as participation in advanced course-taking. Finally, we could consider planning, that is, plans for a future career in a science field. This, of course, would require aspirations for a college degree in a science-related field. Exploring, understanding, and accounting for all 3 aspects of achievement is critical if we hope to cultivate a generation of students who are ready, willing, and able to successfully participate in the science workforce.

The outline of this report is as follows: (1) a review of the empirical literature on gender and race gaps for the 3 Ps, (2) a brief description of our methods (a thorough description is included in the technical appendix of this report), (3) the results, including our descriptive, statistical, and quasi-experimental analyses, and (4) a discussion, including recommendations for NSHS-based staff to consider based on our findings.

Literature Review

In order to provide conceptual clarity for this review of the literature, we focus on three aspects of science education: performance, participation, and planning. These three aspects are interrelated, but analytically, it is worth distinguishing between them. Below, we review the empirical research on each of these aspects, focusing on research that attends to gender differences followed by racial differences.

Performance

A distinction between the domain of performance is necessary to understand the landscape of performance in secondary schools. First, we could consider performance in terms of science assessments (e.g. NAEP, TIMSS, or state- or district-level assessments). Second, we could consider performance in terms of performance in science courses in school (e.g. GPA in science courses).

Looking at performance on assessments first, there is evidence that boys perform better on standardized science assessments than girls (Xie et al., 2015). In a recent analysis of performance on *ECLS-K:99* standardized science assessment, administered as a part of the nationally representative *Early Childhood Longitudinal Study*, Quinn and Cooc (2015) found a small, but significant gap between the scores of eighth grade boys and girls, with boys performing on average 0.185 standard deviations better. However, they also determined that most of this gap is explained by differences in fifth grade math scores. In contrast, girls tend to have better achievement in science courses, with better GPA than boys (Xie et al., 2015).

Similar analyses have related differences in performance by race/ethnicity. While underrepresented minorities have made large strides toward narrowing the achievement gap, they still underperform in relation to White and Asian students (Chen & Soldner, 2014). Looking at achievement on assessments first, there is evidence that there are substantial Black-White and Hispanic-White achievement gaps in science. In addition to the gender gap observed in the Quinn and Cooc (2015) study referenced above, they also found significant Black-White and Hispanic-White gaps in the third grade, with Black students performing on average 1.07 standard deviations lower, Hispanic students .85 standards deviations lower, and Asian students .31 standard deviations lower. By the eighth grade, Asian students had closed the gap and Hispanic students had made significant gains toward narrowing the gap (.20 SD narrower). The Black-White achievement gap, on the other hand, did not change significantly from grade 3 to grade 8. Unlike the gender gap, controlling for prior math achievement substantially reduced, but didn't eliminate, the Black-White and Hispanic-White gaps (.51 and .34 SD lower respectively). Also in contrast to the gender gap, these trends persist in course grades, with White and Asian students typically earning significantly higher grades than Black and Hispanic students throughout all the school years (Kao & Thompson, 2003).

In addition to test scores and course grades, differences in science course taking by race/ethnicity has important implications for college majors and careers beyond graduation. At the K-12 level, minority students tend to take fewer and less challenging math and science courses than their White and Asian peers (Kelly, 2009; Nord et al. 2011). Minority students are also more likely to be placed into remedial and/or low-track math and science courses (Kao & Thompson, 2003). These remedial courses reinforce the course taking gaps we see by race/ethnicity because they tend to cover basic knowledge and involve a lot of rote memorization, which prevents these minority students from ever learning the necessary content and practices to be prepared for higher-level math and sciences courses (Kao & Thompson, 2003).

Participation

In addition to test scores and GPA, course taking in science is critical to understanding the landscape of science education. Advanced course taking prepares students to enter into science-related major, placing them on a trajectory to careers in the sciences (Wang, 2013). Similar, but more notable, gender gaps are observed in course taking. Girls lag behind boys in advance physics course taking in high school, even as they have reached parity in advanced mathematics courses (Riegle-Crumb & Moore, 2013). The gender gap has remained consistent over the past 30 years, even as the number of students taking advanced science courses has increased overall (Riegle-Crumb & Moore, 2013; Xie et al., 2015). In contrast, girls are overrepresented in advanced biology courses, again reflecting a difference in gender gaps by science domain (Xie et al., 2015).

Turning to differences in science course taking by race/ethnicity, we see that at the K-12 level, Black and Hispanic students tend to take fewer and less challenging math and science courses than their White and Asian peers (Kelly, 2009; Nord et al. 2011). Black and Hispanic students are also more likely to be placed into remedial and/or low-track math and science courses (Kao & Thompson, 2003). This overrepresentation of Black and Hispanic students in remedial courses reinforces the course taking gaps we see by race/ethnicity because these courses traditionally focus on basic knowledge and rote memorization. This, in turn, prevents these students from ever learning the necessary content and practices to be prepared for higher-level math and sciences courses (Kao & Thompson, 2003).

Planning

Finally, performance (either on achievement tests or in coursework) and participation in advanced science coursework are for naught if students do not then aspire to enter a science-related major and/or career field. We call this planning. Planning probably interacts with performance, affect, and interest (Xie et al., 2015). Research exploring planning often looks at intention to declare a science major among high school students, desire for a career in the sciences, or interest in science as a field. Research also exists that explores trends in declared majors and careers, which occurs after secondary schooling and so is slightly outside the scope of this paper. However, since planning is key to the story, we touch on some of the empirical literature.

Gender gaps in planning are similar to gaps in performance, but more pronounced and appear more intractable. While girls are doing better in science (i.e., GPA), they are not aspiring to science career paths to a greater extent. As Xie and colleagues (2015) succinctly note, preparation has not translated to participation for women. Looking at STEM majors more generally, Wang (2013) found that 11% of girls intend to major in a STEM field, compared to 29% of boys. Analyzing nationally representative surveys, Riegle-Crumb and colleagues (2012), found that there is an even more pronounced gender gap in declared majors when looking within science domain. While women and men declared majors in biological sciences at similar rates, with about 10% of women and men declaring, there was a pronounced gender gap in the declaration of a major in the physical sciences or engineering. About 25% of men declared a major in the physical sciences, compared to about 5% of women.

Similar to the physical sciences versus natural sciences gaps in declared majors, girls express intentions for careers in the natural sciences at a much greater rate than in the physical sciences (Riegle-Crumb et al., 2010). These aspirations vary by race/ethnicity, with greater parity among White boys and girls than Hispanic or Black. Riegle-Crumb and colleagues (2012) also explored differences in science enjoyment and self-concept—how able students believe they are in the subject. In general, boys showed higher average enjoyment and self-concept in science than girls. However, again, the differences varied by race/ethnicity, with a smaller difference between white boys and girls and larger differences between Hispanic and Black boys and girls. Other research has demonstrated gaps in interest and affect between boys and girls, showing a persistent difference between the genders. That is, even when controlling for achievement, girls show less interest, motivation, self-confidence, and desire for a career in STEM fields (Xie et al., 2015).

While the share of minorities entering college and obtaining a STEM degree has increased over time, they are still underrepresented. For example, the share of minorities among all recipients of science and engineering bachelor's degrees grew from 17% in 2000 to 20% in 2011. Still, minority students are underrepresented because their share in the general population aged 25–29 is much higher, above 36% in 2011.

Method

Our analyses centered around one main question of interest: What variables can predict student achievement in science? As mentioned above, the literature suggests that "achievement" can mean many things. In our analyses, we focused on three aspects of science education: performance, participation, and planning (see Table 1). The performance variables we used are those that are closely related to traditional conceptions of achievement, including GPA and ACT scores. The participation variables we used are course-taking decisions by the students including participation in AP classes. We also looked at the number of science courses taken above the NSHS requirement (i.e., 4 science courses). Finally, the planning variables we used are students' post-high school graduation plans including intended college major(s). These variables came from the senior survey.

Outcomes of Interest	**Variables**
Performance	-Science ACT -Cumulative science GPA
Participation	-Number of science courses taken (above requirement) -AP science course-taking
Planning	-Intended science major

Table 1 - Plan for analysis, dependent and independent variables

Our analysis took three main parts. First, we conducted a descriptive overview of achievement in science at NSHS – documenting a descriptive overview for all three aspects of achievement. In particular, we wanted to investigate whether certain students are disproportionately likely to struggle. That is, we found the mean and standard deviation of all our variables of interest by our demographic variables (including gender, race/ethnicity, and free and/or reduced price lunch). While not causal, this descriptive analysis provided an important overview to the extent to which students at NSHS are succeeding in science, who is succeeding, and when they are succeeding.

Along with a descriptive picture of science achievement at NSHS, we also conducted a statistical analysis. While the descriptive analysis provided a snapshot of science achievement at NSHS, this snapshot encompassed a lot of factors that contributed to the outcomes we examined, including ones that are already present when students arrived in their freshman year. This includes income, what they learned in middle school, etc. The statistical analysis allowed us to rule out a number of things that caused an achievement gap between students before they ever set foot in NSHS. This happened by matching students who are as similar as possible on the variables we had access to. For example, the statistical analysis compared science ACT scores for two students who were both from high-income families because we know that income plays a role in achievement (this is for a multitude of reasons that we won't get into here).

Although our statistical analysis tells us more about the variables that contribute to student achievement in science, these are associations. That is, people who tend to do X, tend to do X. However, correlation does not equal causation so even our statistical analysis can't tell you with certainty that if you changed X, Y would change accordingly. Therefore, for our final analysis, we conducted a quasi-experimental analysis that explored plausible causal effects of NSHS's science placement policy for freshman honors biology. Since NSHS placed students into either honors or regular biology based primarily on middle school EXPLORE reading test scores, we can compare students just above the threshold for placement in honors biology with those just below the threshold. The idea is that these groups of students are very similar in many ways relevant for science achievement. Using this logic, those just above the cutoff are the "treatment group", who receive honors biology their freshman year, and those just below are the "control group", who did not receive honors biology. By comparing these two groups, we are able to determine the impact of taking an honors biology class freshman year on the outcomes of interest. Making the comparison occurs in two steps. First, we find the likelihood of taking freshman honors biology based on EXPLORE reading scores. Second, we use that likelihood to determine the impact of probability of taking freshman honors biology on the three Ps. A more detailed explanation of this method is given in the technical appendix.

Results

To identify the factors that are associated with science achievement, we examined how demographic variables predicted student outcomes. Below, we present the results from our main analysis of three performance outcomes. First, we explore the relationship between participation in science courses above those required and in AP science classes, and a set of demographic variables, focusing specifically on differences by gender and ethnicity. We then look at the relationship between demographics and traditional performance outcomes (GPA for science classes and the science portion of the ACT assessment). Last, we explore the relationship between demographics and explicit plans to major in a science-related field in college. In each analysis, we focus on the differences in performance by gender and by ethnicity. We produce an overall analysis for the relationship between science performance outcomes and demographics, and then look within each gender and ethnicity to identify key differences. By doing this, we can explore if science course taking, for example, has a differential impact for boys or girls, or for White or Hispanic students. Looking across ethnicities, we can also see if there are significant gender gaps within these subgroups. Looking across genders, we can identify significant ethnic gaps within each gender.

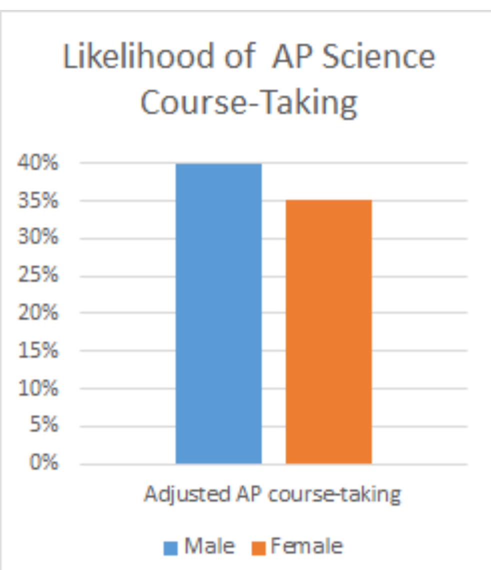
Participation

Before examining cumulative GPA or ACT scores which are measured at the end of a student's time at NSHS, we want to understand how differences in performance outcomes might occur - that is through a commitment to science course-taking. We present the results from our main analysis of two participation outcomes in this section. First, we explore the relationship between the total number of science courses taken above the four that are required for NSHS graduation and a set of student demographic variables, focusing on specifically on subgroup differences by gender and ethnicity. We then look at the relationship between these demographics and AP course-taking. These analyses provide a view into which students, on average, pursue a science and college preparedness track.

Gender differences.

Consistent with literature that highlights a gender gap in pursuing science majors and related careers, girls took 0.21 fewer additional science courses than did boys, net of other characteristics such as ethnicity, past performance, and income status. Family income, Middle School Explore scores in Science, Math and English and cumulative science GPA were all significant predictors of participation in additional science courses, with higher Explore scores and science GPA corresponding with taking more classes, while low-income corresponded with taking fewer.

There was also an overall gender gap in AP science participation as was found with science course-taking more broadly, such that girls were 4.6% less likely to take an AP class in a science subject, net of other characteristics such as ethnicity, past performance, and income status.



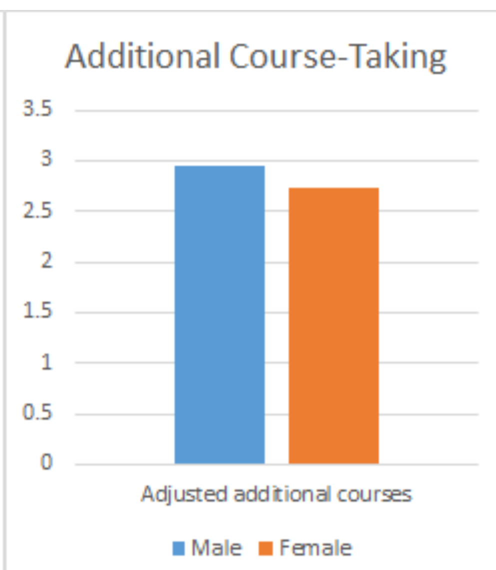


Figure 1. Science participation by gender, adjusted for previous performance and other factors.

Racial differences.

With additional course-taking in science, we see a pattern that arises in previous research: white and asian students are taking more science courses than black & hispanic students. After adjusting for other factors though, the gap shrinks. Asian students took 0.2 more science courses than were required, net of other characteristics, but no other ethnicity significantly predicted additional course-taking. Notably, participation in a free and reduced-price lunch program predicted decreases in course-taking for white students only, suggesting that a general tendency for this population to take more science courses might correspond with greater family resources.

Race and ethnicity were significant predictors for the likelihood of AP science course-taking for Black and Hispanic students, net of other characteristics such as gender, past performance, and income status. Black students were 6.2% less likely to take these classes than their White counterparts and Hispanic students were 7% less likely to do so.

Science GPA was a predictor of increases in AP course-taking, though this may be mechnical if certain grades in pre-requisites are required to take these AP courses. Non-science GPA was a predictor for increases AP course-taking across all ethnic groups except for Asian students. For taking AP classes, low-income was a predictor only for Multi-racial students, with participation in free or reduced price lunch correlating with 15.5% decrease in taking AP courses for this demographic.

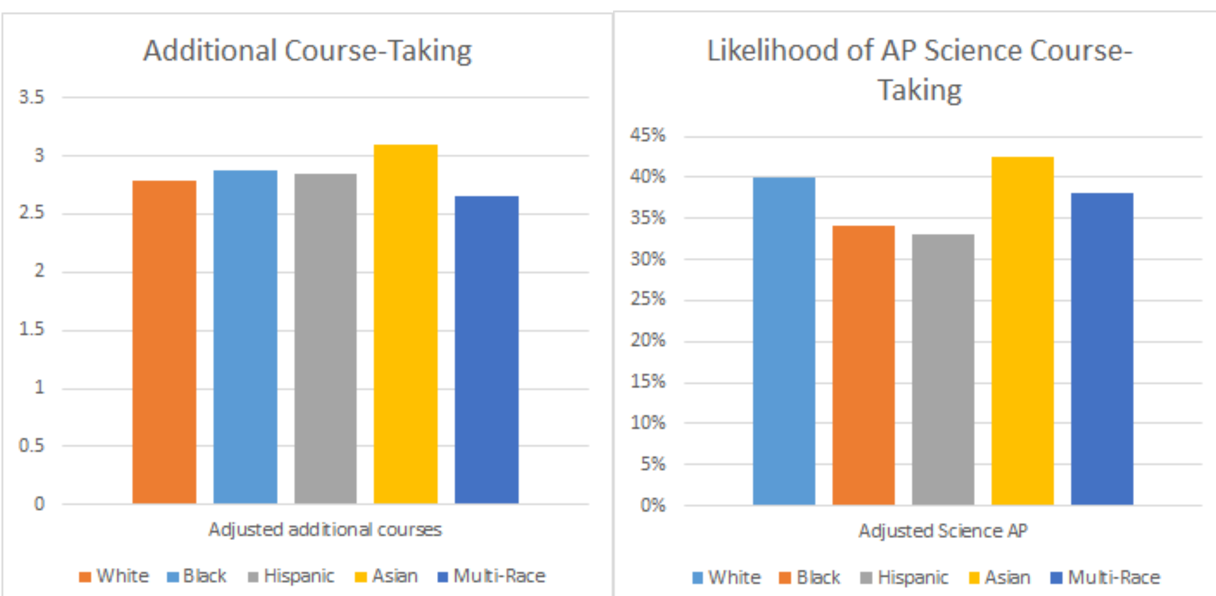


Figure 2. Science participation by race and ethnicity, adjusted for previous performance and other factors.

Racial & gender differences.

While many of the relationships mentioned hold across all groups, there are notable exceptions. There is a persistent gender gap in additional course-taking across most ethnic groups, with the smallest gap for white students and the largest for multi-racial students. Sex was not a significant predictor of additional course-taking for Hispanic students.

While girls were, in general, less likely to take an AP science course, this effect was not present in all ethnicities. Sex did not predict AP course-taking for Asian students and was not a significant predictor for White or Black students. Still, female Hispanic students were 6.7% less likely to take AP science courses than their male counterparts and female multi-racial students were 15.5% less likely to participate in these courses, net of other characteristics.

Performance

Below, we present the results from our main analysis of two performance outcomes. First, we explore the relationship between performance on the science portion of the ACT assessment and a set of demographic variables, focusing specifically on differences by gender and ethnicity. We then look at the relationship between demographics and science GPA.

Some general findings are that past performance on EXPLORE tests, science and non-science GPA, and total number of science courses taken all predict performance on the science ACT. Each of these has an independent impact on science ACT scores. For example, regardless of prior performance (as measured by EXPLORE assessments), students on average get a 0.238 bump on their science ACT for each additional science course that they take. We should note here that this does not show a causal relationship between course taking and science ACT scores. There may exist an unobserved variable that is correlated with both increased science course taking and science ACT scores, such as interest in science. Notably, science and non-science GPA are the strongest predictors, suggesting that those student who do well in class (regardless of gender, ethnicity, or past performance) have better science ACT scores.

For science GPA, only performance on the science EXPLORE test and not the reading or math test is predictive. There is a strong significant relationship between non-science GPA and science GPA, showing that students who do well in non-science subjects also tend to do well in science. In contrast to science ACT, there is a weak, marginally significant relationship between science course taking and science GPA. A few reasons may explain this. For example, as student take more science courses, they move into more advanced subject matter, and therefore get lower grades. Notable across all models, there is no difference in science GPA by income status, accounting for the other predictors in the models.

Gender differences.

Overall, boys perform about 1.2 points better on the ACT, net of other characteristics such as ethnicity, past performance, science GPA, course-taking, and income status. Girls also have, on average, a lower science GPA than boys, net of other characteristics. Notably, taking simple averages, girls have higher science GPAs (M= 3.21 vs M= 3.09 for boys) and higher overall GPAs. However, once we control for the other predictor variables in our model, we find that girls have worse science GPAs, and when non-science GPA is taken into account, the coefficient on the model is negative and significant. This means that girls perform worse in science courses than boys who have similar levels of achievement in other classes, suggesting that a gender gap in science is growing during students' time at NSHS.

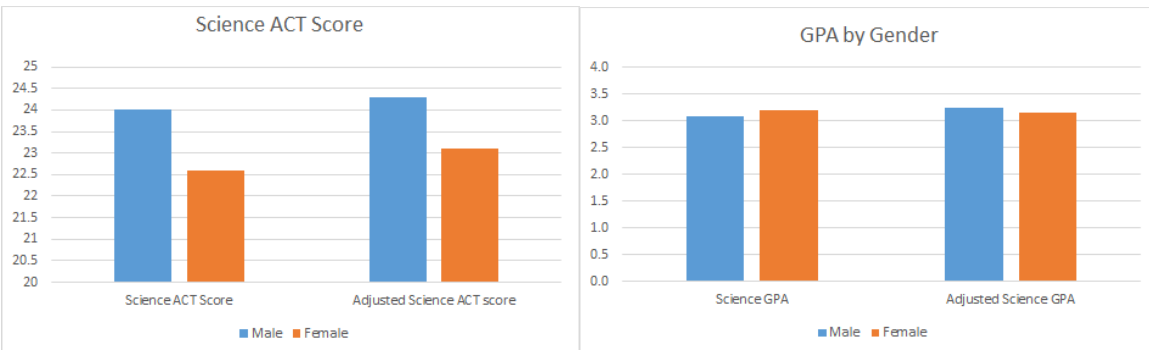


Figure 3. Science performance by gender, before and after adjusting for previous performance and other factors.

Racial/ethnic differences.

Black and Hispanic students perform 1.2 and 1 points worse on the science ACT than White students, net of other characteristics. For cumulative science GPA, we again see Black and Hispanic students have on average lower science GPAs than White students. This pattern mirrors AP course-taking, perhaps unsurprising given the weighted GPA that comes with taking the competitive course tracks, but these findings are net of other predictors in the model, so the GPA gap is not due to worse prior performance, income status, the number of science courses taken, or performance in other courses.

For Black students there is a smaller impact on science ACT from prior science performance and GPA that the overall model. However, taking courses in science is more beneficial for Black students than White students. The pattern is similar for Hispanic students, except that GPA in science classes and not other classes impacts their ACT scores.

Looking across the major ethnic groups represented at NSHS, we see that non-science GPA is strongly associated with science GPA, suggesting that students who do well in other classes do well in science. Black and Hispanic students also see a boost in science GPA from prior science performance (i.e., science EXPLORE scores). The relationship is strongest for Hispanic students, where the impact is almost double that of the main specification and double the impact seen by Black and White students. The impact of course-taking also varies across ethnicities. Asian and White students see a larger boost from each additional science course they take on their overall science GPA. For Asian students, the impact is over six times larger than the main specification. For Black students, the relationship is negative. As Black students take more science courses, they see a decrease in science GPA. Since many factors contribute to course grades, it is difficult to speculate on what might be driving these differential impacts.

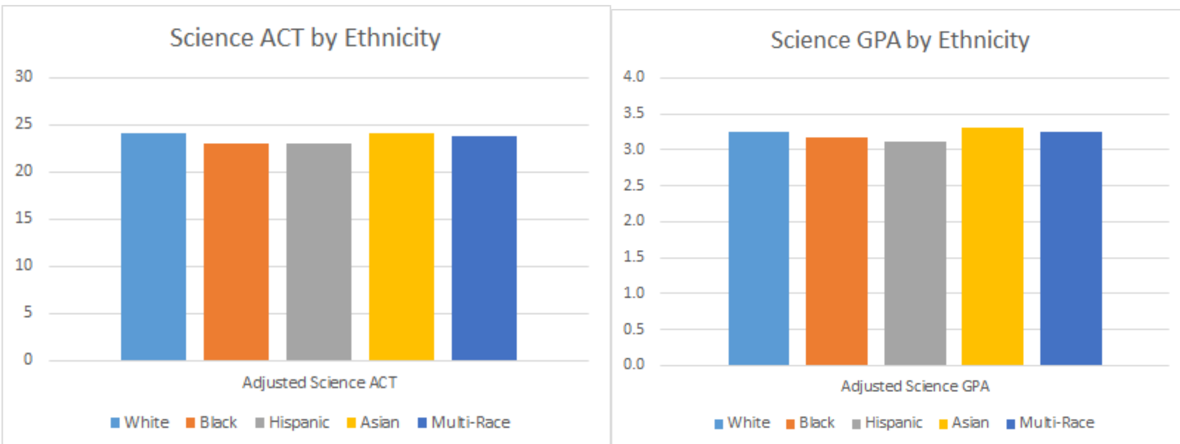


Figure 4. Science participation by race and ethnicity, adjusted for previous performance and other factors.

Racial and gender differences.

For ACT scores on the science portion, there is a persistent gender gap across ethnic groups, with the smallest gap for Black students and the largest for multi-racial students. Black and Hispanic students have lower science ACT scores on average than White students, but the gap is larger among girls than among boys. The gap between the performance of girls and boys is also larger for Asian students than in the main specification; only the gap for multi-racial students is larger. Female Asian students have science ACT scores that are 1.6 point lower than boys.

We see a similar impact on science GPA from prior science performance and from non-science GPA for both boys and girls. The main differences relate to ethnic gaps in science GPA. The ethnic gaps are larger for girls that for boys. For boys, there is no gap between White students and Black students, although a gap still remains between White and Hispanic students. For girls, the gaps between Black or Hispanic and white students is larger than in the main specification. Hispanic and Black girls have, on average, worse science GPAs that White girls. Again, this gap persists even when taking into account prior performance, GPA in other courses, and the number of science courses taken.

Planning

To identify the factors that are associated with science planning, we explored the relationship between students' intent to major in a science-related domain in college and a set of demographic variables. Intended majors were determined by free response items in the senior survey and were coded as 0 for majors not in a traditional science area or 1 for science majors (e.g., chemistry, physics, biology, environmental science). Also coded as 1 were applied science paths like engineering (i.e., any major with engineering in the name) and medicine (e.g., pre-med, veterinary). Nursing and radiology were not coded as science majors, however if students had listed a traditional science major along with an applied (e.g., "pre-med, biology") the intended major was coded as 1.

Gender differences.

Figure 5 shows that when we control for gaps that we see upon arrival at NSHS, as well as GPA and number of science courses taken, boys have a 2.9 percentage points greater chance of intending to major in science in college. However, this difference is only marginally significant.

Racial differences.

This is encouraging because it is a much smaller gap than we found in other outcomes of interest. It is also encouraging to find that with one exception there are no meaningful differences between racial/ethnic groups once we control for observable covariates (see figure 5). The exception is that multi-race students are 10.7 percentage points more likely to intend on a science major in college. It's particularly difficult to speculate about what is driving this effect because these students represent a group of students with a wide range of experiences.

Racial & gender differences.

While there were no statistically significant differences between non-White and White students chances of intending to major in science when looking within ethnicity (with the exception of multi-race), within gender, Asian girls are more likely to intend on a science major. Additionally, the primary driver in the gap between multi-race students and white students seems to be due to male multi-race students.

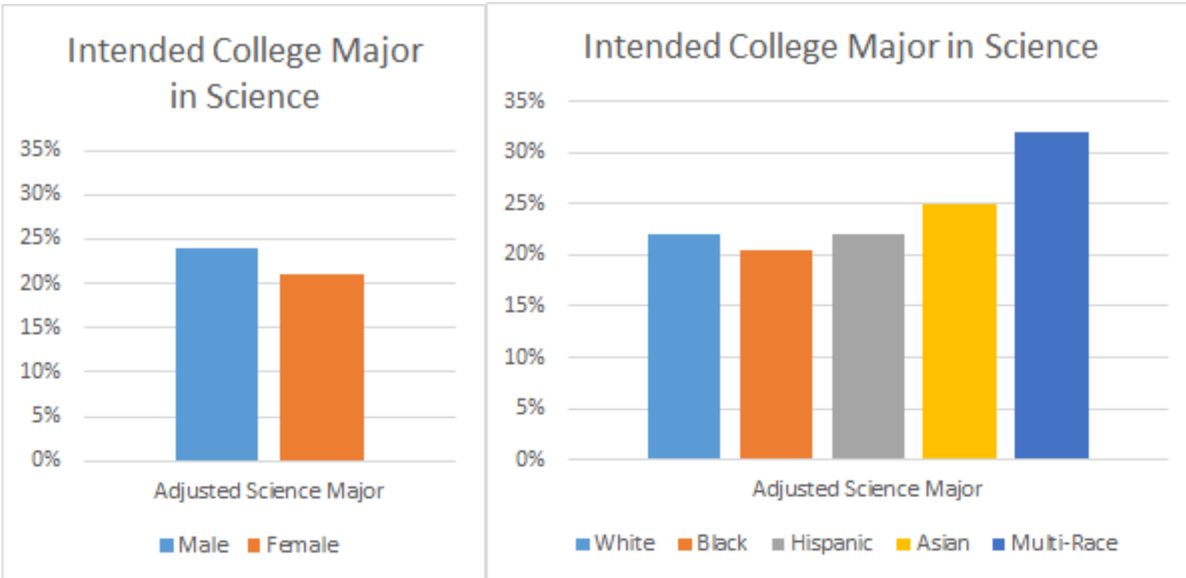


Figure 5. Stated intent to major in a science-related field in college, comparing by gender and by race and ethnicity, adjusted for previous performance and other factors.

In addition to our finding gender and racial/ethnic group gaps (or lack thereof in this case) in the intent to major in a science-related domain in college, we found a number of other interesting predictors. Not surprisingly, the total number of science courses taken in high school was a strong predictor. Science and non-science GPAs were also strong predictors, but in opposite directions. For every additional 1.0 increase in science GPA, students, on average, have a 12.4 percentage point increase in the chance they intend to major in science in college. On the other hand,, for every additional 1.0 increase in non-science GPA, students, on average, have a 13.1 percentage point decrease in the chance they intend to major in science in college. This suggests that students are aware of what they are good at and want to further pursue knowledge domains that they were successful with in high school. It also suggests that some talented students are not going into science because they feel they are relatively worse or it or at least better in other domains.

Evaluation of placement in freshman honors biology

The preceding analyses of the relationship between student characteristics and science achievement offer statements of associations, but not causal relationships. Moreover, the analyses focus heavily, but not exclusively, on characteristics of students the moment they enter NSHS, which school policy at NSHS cannot address. For these reasons, we decided to evaluate the impact of freshman honors biology at NSHS, exploiting the fact that, until recently, placement into honors biology was determined in large part by middle school EXPLORE test scores. This analysis allows us to focus on a specific NSHS policy, as well as offer a causal evaluation of the policy.

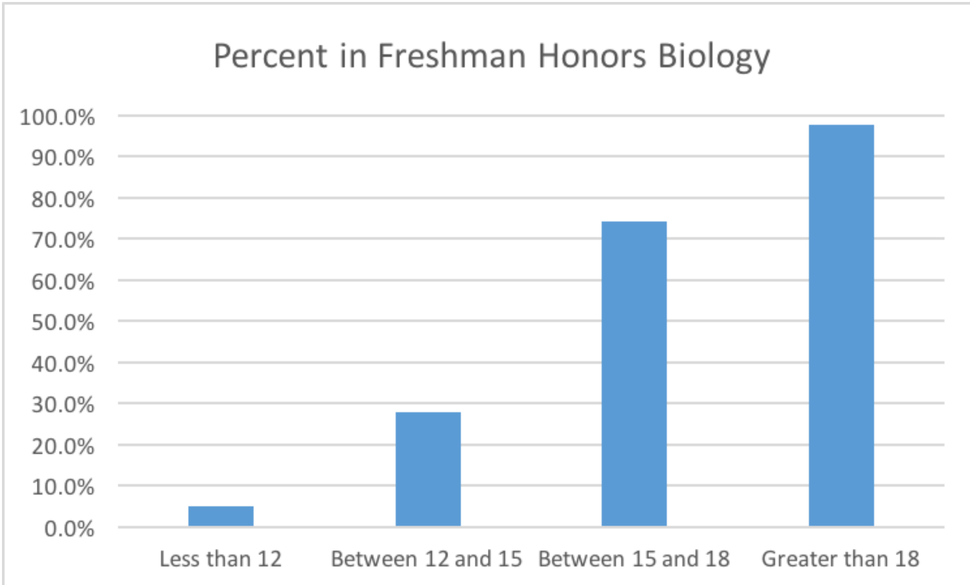


Figure 6. Percent taking honor biology freshman year by reading EXPLORE score

Students who score 15 or above on their reading EXPLORE scores in middle schools were eligible to enroll in honors biology in their freshman year. While this was not a hard-and-fast rule--some students below the cutoff took honors biology, some students above the cutoff did not--students who were just above the cutoff were much more likely to enroll in honors biology than those just below the cutoff. Of the students just above the cutoff, 74% enrolled in honors biology, compared to just 28% of students just below the line (Figure X). Our analysis of placement into freshman honors biology compares the outcomes of students with EXPLORE reading scores between 15 and 17 (inclusive) who qualified for honors biology to those between 12 and 14 (inclusive). We compare these two groups because they are likely very similar on many dimensions. They have similar reading ability and likely similar motivation. The key difference is that those just above the cutoff are able to take honors biology, while those just below take regular biology. The first group are the "treatment" group, while the latter are the "control" group. This allows us to plausibly assess the causal impact of honors biology on science achievement. More details about our analytic strategy and estimation results are available in the Technical Appendix.

As before, we explore the impact of honors biology on the three Ps of science achievement. In each case, except one, we find that students placed in honors biology perform no better or worse than very similar students placed in regular biology. We cannot detect an impact on science ACT, science course taking, or the likelihood to intend a science major. We find an impact on science GPA; however, this is at least partly due to the additional credits of honors and AP course. When we run the analysis using an unweighted GPA (i.e. an A in an honors class is 4 points, rather than 4.5 points), the effect goes away.

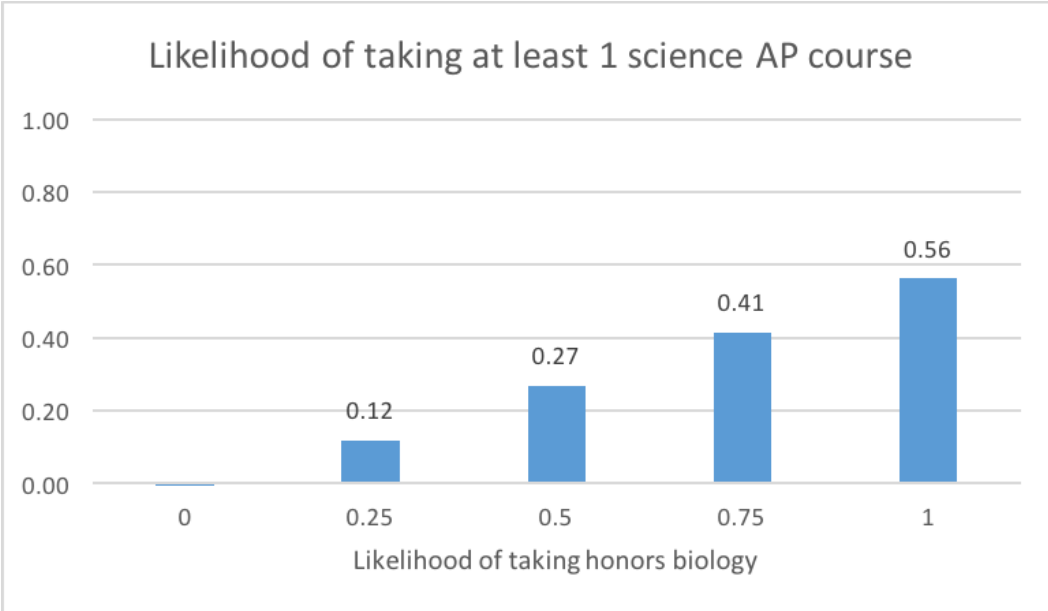


Figure 7. The likelihood of taking at least one science AP course based on the likelihood of taking honors biology freshman year

Thus, the impact of honor biology on GPA is partly mechanical.

Importantly, we find a modest, but important, impact on the likelihood of taking at least on science AP course. Compared to students just below the cutoff, students just above the cutoff who take honors biology are on average 56 percentage points more likely to take honors biology. As the likelihood that a student takes freshman honors biology increase, so does the likelihood that she takes at least one AP course. Again, this result compares very similar students who had similar scores on their middle school EXPLORE reading tests. Thus, by taking honors biology in their freshman year, these students are more likely to take more advanced science courses.

Discussion

To summarize, gender and race or ethnicity predict differences in science achievement across participation in science courses and performance in classes and the ACT, though not in the intention to major in a science-related area, such that male students demonstrate higher achievement than female students and White and Asian students have higher achievement than Black and Hispanic students. When we matched students on similar characteristics, the gaps in science achievement that are apparent when students arrive at NSHS tend to be reduced during a student's time in high school. However, these trends are not reversed. Even among similarly performing students, Black & Hispanic students tend to suffer across science achievement outcomes, though to a lesser degree. And between girls and boys in science, the gaps in participation and performance are either persistent or even increase over the course of high school, as is the case with science GPA, despite similar performance in non-science GPA. What can NSHS do with this information?

Based on our regression discontinuity results, the placement of average-performing students into Honors Biology appears to make them more likely to take at least one AP class in the future, opening doors to higher education. To promote this, NSHS may consider expanding honors access in this introductory science class. Recent efforts promoting earned honors may thus increase AP course-taking in science, though future research will need to determine whether the effects stem from the weighted grades, a cohort of students interested in pursuing science in college, or other factors.

Additionally, while the achievement gaps between groups are reduced when we consider students who begin at similar academic levels, many of the gaps we document emerge before students enter NSHS, contributing to their academic paths and outcomes. Therefore, NSHS may also consider ways to provide support for students entering NSHS at different levels. Efforts at reducing these differences should be targeted to students at all levels (for example, even high achieving girls have lower achievement than boys).

Other recommendations are for facilitating future research. With the current data available, it was difficult to estimate the effect of honors placement alone given that Reading Explore scores were a fairly reliable, but not definitive placement criteria.
To better understand the impact of placement in honors or other programs, NSHS could explore whether there ways to clarify placement criteria. For example, NSHS could formalizing teacher references on a scale of 1 to 5, including these in student records. This would allow researchers in the future to do a more precise job of identifying and comparing similar students.

References

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Committee on Science, Engineering, and Public Policy. (2007). Rising above the gathering storm: Energizing and employing America for a brighter economic future. National Academies.

Muller, P. A., Stage, F. K., & Kinzie, J. (2001). Science Achievement Growth Trajectories: Understanding Factors Related to Gender and Racial-Ethnic Differences in Precollege Science Achievement. *American Educational Research Journal*, *38*(4), 981–1012.

Quinn, D. M., & Cooc, N. (2015). Science Achievement Gaps by Gender and Race/Ethnicity in Elementary and Middle School: Trends and Predictors. *Educational Researcher*, *44*(6), 336–346.

Riegle-Crumb, C., & King, B. (2011). Questioning a White Male Advantage in STEM: Examining Disparities in College Major by Gender and Race/Ethnicity. *Educational Researcher*, *39*(9), 656–664.

Riegle-Crumb, C., & Moore, C. (2013). The Gender Gap in High School Physics: Considering the Context of Local Communities. *Social Science Quarterly*, *95*(1), 253–268.

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The More Things Change, the More They Stay the Same? Prior Achievement Fails to Explain Gender Inequality in Entry Into STEM College Majors Over Time. *American Educational Research Journal*, *49*(6), 1048–1073.

Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2010). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education*, *95*(3), 458–476.

Wang, X. (2013). Why Students Choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support. *American Educational Research Journal*, *50*(5), 1081–1121.

Xie, Y., Fang, M., & Shauman, K. (2015). STEM Education. *Annual Review of Sociology*, *41*(1), 331–357.

Technical Appendix

For the main analysis of our paper, we conducted a series of regression analyses for each of our outcomes of interest. We chose a series of demographic and achievement variables available to us that are known in the literature to related to science achievement. We conducted a main analysis for each outcome, along with sub-analyses for gender and ethnicity. These latter are "fully interacted" models that capture the differential impact of predictors for each sub-group. Thus, we can provide a picture of the overall relationship between our predictors and science achievement, and we show the relative importance of each predictor for each subgroup. It is important to note that in the absence of a quasi-experimental design these analyses should be viewed as providing suggestive evidence of relationships, rather than causal estimates.

Exploiting NSHS's previously used policy for placing students into honors Biology in freshman year, we designed a quasi-experimental analysis using regression discontinuity. This approach allows for plausible causal estimates of the impact of policies on student outcomes. In this case, we explore how participation in honors Biology freshman year impacts each of the three Ps of our main analysis. The advantage of this approach is that it gives greater credence to causal statements. An important limitation is that it identifies causal impacts only for a small group of students--those just above and just below the cutoff for placement into honors Biology.

Regression Analysis

Our regression analysis plan for each set of outcomes follows the same procedure. First, we define a main specification, including gender, ethnicity, and all the relevant covariates. Second, we conduct analyses of subgroups by restricting the regression to include only members of each gender and ethnic group. In the case of gender, the indicator for gender is left out, while all other covariates are included. In the case of ethnicity, the indicator for ethnicity is left out, while all other covariates are included. The subgroup analyses equivalent to fully interacted models, in which the subgroup of interest is interacted with all other terms in the model. In practice, this was achieved by restricting regression to include only members of the subgroup (e.g. only female students). In what follows, we define our models for each outcome and present regression results.

Performance

To assess achievement via academic performance, we perform a series of OLS regressions in which the outcomes of interest are the students' science-specific GPA and ACT science scores (all NSHS students are required to take the ACT). Our main specification is year fixed-effects OLS model:

$$(1)\ A_{ic} = \alpha + \beta_1 E_{ic} + \beta_3 F_{ic} + \mathbf{X}_{ic}\beta + \delta_c + \varepsilon$$

Where A_{ic} is the performance outcome for student *i* in cohort *c*; E_{ic} is an indicator variable for the race of student *i* in cohort *c*; F_{ic} is a dummy for whether student *i* in cohort *c* is female; $\mathbf{X}_{ic}$ is a vector of student characteristics; and δ_c is a cohort fixed-effects term. The vector of student characteristics includes prior achievement (8th grade explore scores in math, science, and reading), science GPA, non-science GPA, free/reduced price lunch status, and total number of science courses taken. We run the main specification and then individual specifications for subgroups by race and gender to capture within subgroup associations. When we run the regression of science GPA on race and gender, we leave science GPA out of the regression. When we run the regression of AP scores on race and gender, we use science GPA from freshman and sophomore years, since taking an AP class gives a GPA 'bonus.'

Tables A1 and A2 show the results for performance outcomes by gender. The first column shows the main specification, followed by the specification restricted to boys and the specification restricted to girls. Tables A3 and A4 show the results for performance outcomes by ethnicity.

Table A1 - Results for Science ACT by gender

VARIABLES	(1) Main Specification	(2) Within boys	(3) Within girls
Female	-1.208***		
	(0.117)		
Black	-1.170***	-0.715**	-1.713***
	(0.169)	(0.241)	(0.233)
Hispanic	-1.040***	-0.562*	-1.609***
	(0.197)	(0.275)	(0.276)
Asian	0.0446	0.296	-0.179
	(0.289)	(0.398)	(0.404)
Multi-race	-0.443	-0.378	-0.619
	(0.280)	(0.377)	(0.400)
Cumulative Science GPA	0.812***	1.166***	0.548*
	(0.155)	(0.223)	(0.214)
Cumulative non-Science GPA	0.959***	0.650*	1.184***
	(0.218)	(0.298)	(0.317)
EXPLORE Reading	0.188***	0.162***	0.213***
	(0.0227)	(0.0323)	(0.0317)
EXPLORE Science	0.312***	0.314***	0.311***
	(0.0320)	(0.0436)	(0.0459)
EXPLORE Math	0.286***	0.302***	0.254***
	(0.0247)	(0.0359)	(0.0342)
EXPLORE English	0.143***	0.203***	0.0855**
	(0.0218)	(0.0316)	(0.0302)
Total number of science courses	0.286***	0.296***	0.285***
	(0.0466)	(0.0675)	(0.0651)
Free/reduced price lunch	-0.108	-0.437	0.203
	(0.158)	(0.228)	(0.214)
Constant	0.0321	-1.125	0.136
	(0.543)	(0.753)	(0.785)
Observations	3,054	1,490	1,564
R-squared	0.750	0.763	0.737
Cohort FE	YES	YES	YES

Robust standard errors in parentheses
*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table A2 - Results for Science cumulative GPA by gender

VARIABLES	(1) Main Specification	(2) Within boys	(3) Within girls
Female	-0.115***		
	(0.0143)		
Black	-0.0962***	-0.0448	-0.142***
	(0.0224)	(0.0314)	(0.0319)
Hispanic	-0.139***	-0.131***	-0.150***
	(0.0248)	(0.0358)	(0.0347)
Asian	0.00122	0.0129	-0.0223
	(0.0258)	(0.0423)	(0.0327)
multi-race	-0.0388	-0.0323	-0.0501
	(0.0342)	(0.0391)	(0.0588)
Cumulative non-Science GPA	1.185***	1.161***	1.217***
	(0.0156)	(0.0191)	(0.0253)
EXPLORE Reading	-0.000635	-0.00385	0.00280
	(0.00243)	(0.00347)	(0.00340)
EXPLORE Science	0.0158***	0.0168***	0.0142**
	(0.00368)	(0.00509)	(0.00523)
EXPLORE Math	-0.000614	0.00204	-0.00158
	(0.00293)	(0.00390)	(0.00436)
EXPLORE English	-0.000976	-0.000445	-0.00251
	(0.00254)	(0.00346)	(0.00365)
Total number of science courses	0.0124	0.0185	0.00369
	(0.00653)	(0.00958)	(0.00892)
Free/reduced price lunch	-0.00675	0.00801	-0.0123
	(0.0203)	(0.0295)	(0.0285)
Constant	-1.055***	-1.058***	-1.188***
	(0.0683)	(0.0938)	(0.0985)
Observations	3,285	1,621	1,664
R-squared	0.875	0.872	0.879
Cohort FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table A3 - Results for Science ACT by ethnicity

VARIABLES	(1) Main Specification	(2) Within White	(3) Within Black	(4) Within Hispanic	(5) Within Asian	(6) Within multi-race
Black	-1.170***					
	(0.169)					
Hispanic	-1.040***					
	(0.197)					
Asian	0.0446					
	(0.289)					
multi-race	-0.443					
	(0.280)					
Cumulative Science GPA	0.812***	1.165***	0.629**	1.017**	-0.436	0.406
	(0.155)	(0.288)	(0.225)	(0.349)	(1.219)	(0.957)
EXPLORE Reading	0.188***	0.155***	0.194***	0.191**	0.322**	0.222
	(0.0227)	(0.0299)	(0.0494)	(0.0673)	(0.0966)	(0.130)
EXPLORE Science	0.312***	0.360***	0.191**	0.168	0.361*	0.364*
	(0.0320)	(0.0420)	(0.0637)	(0.0892)	(0.145)	(0.169)
EXPLORE Math	0.286***	0.284***	0.326***	0.274***	0.300*	0.0359
	(0.0247)	(0.0339)	(0.0507)	(0.0641)	(0.125)	(0.125)
EXPLORE English	0.143***	0.138***	0.174***	0.127*	0.0379	0.301**
	(0.0218)	(0.0299)	(0.0462)	(0.0621)	(0.0943)	(0.103)
Total science courses	0.286***	0.283***	0.338***	0.413**	0.628*	-0.0804
	(0.0466)	(0.0728)	(0.0826)	(0.125)	(0.285)	(0.237)
FRPL	-0.108	-0.145	-0.123	-0.0291	-1.521*	-0.0895
	(0.158)	(0.291)	(0.249)	(0.372)	(0.586)	(0.790)
Female		-1.381***	-0.914***	-1.192***	-1.591**	-2.099***
		(0.166)	(0.229)	(0.288)	(0.562)	(0.603)
Non-Science GPA		0.916*	0.625	0.841	1.635	2.926*
		(0.380)	(0.335)	(0.515)	(1.560)	(1.304)
Constant	0.0321	-1.312	0.762	0.864	-0.874	-2.099
	(0.543)	(0.868)	(0.897)	(1.293)	(3.478)	(2.878)
Observations	3,054	1,603	768	440	115	128
R-squared	0.750	0.597	0.593	0.657	0.746	0.782
Cohort FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses; FRPL = free/reduced price lunch

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

Table A4 - Results for cumulative Science GPA by ethnicity

VARIABLES	(1) Main Specification	(2) Within White	(3) Within Black	(4) Within Hispanic	(5) Within Asian	(6) Within multi-race
Female	-0.115***	-0.050**	-0.231***	-0.128**	-0.054	-0.120
	(0.014)	(0.015)	(0.034)	(0.044)	(0.054)	(0.076)
Black	-0.096***					
	(0.022)					
Hispanic	-0.139***					
	(0.025)					
Asian	0.001					
	(0.026)					
multi-race	-0.039					
	(0.034)					
Non-Science GPA	1.185***	1.153***	1.190***	1.225***	1.133***	1.160***
	(0.016)	(0.021)	(0.028)	(0.042)	(0.078)	(0.086)
EXPLORE Reading	-0.001	-0.003	-0.003	-0.000	0.005	0.020
	(0.002)	(0.003)	(0.007)	(0.009)	(0.010)	(0.013)
EXPLORE Science	0.016***	0.017***	0.016	0.030**	-0.001	-0.012
	(0.004)	(0.004)	(0.009)	(0.012)	(0.015)	(0.018)
EXPLORE Math	-0.001	0.008**	-0.004	-0.008	0.019	-0.018
	(0.003)	(0.003)	(0.007)	(0.011)	(0.010)	(0.013)
EXPLORE English	-0.001	-0.005*	0.009	-0.006	-0.009	0.019
	(0.003)	(0.003)	(0.007)	(0.009)	(0.008)	(0.013)
No. of science courses	0.012	0.055***	-0.041**	-0.006	0.080***	0.011
	(0.007)	(0.007)	(0.013)	(0.019)	(0.022)	(0.043)
FRPL	-0.007	-0.003	0.038	-0.075	-0.066	-0.122
	(0.020)	(0.034)	(0.036)	(0.048)	(0.060)	(0.098)
Constant	-1.055***	-1.351***	-0.839***	-1.180***	-1.182***	-0.980*
	(0.068)	(0.080)	(0.123)	(0.182)	(0.262)	(0.459)
Observations	3,285	1,695	865	469	119	137
R-squared	0.875	0.859	0.745	0.810	0.916	0.848
Cohort FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Participation

Something of particular interest to NSHS stakeholders that we can determine from these regressions is the relationship between course-taking and achievement in science. We will estimate a series of OLS regression models to capture key associations between variables of interest and science course taking patterns. Our main specification is similar to (1):

$$(2)\ P_{ic} = \alpha + \beta_1 E_{ic} + \beta_2 F_{ic} + \mathbf{X}_{ic}\beta + \delta_c + \varepsilon$$

The term P_{ic} refers to our participation outcomes of interest. There are two outcomes we explores: (1) the total number of science courses taken over the four required and (2) the total number of AP classes taken. The remaining terms are as specified above, except for the vector of student characteristics. From that vector, we exclude the total number of science classes taken. We include science GPA from freshman and sophomore years rather than overall cumulative GPA for the regression of AP courses on gender and race. As before, we run the main specification, followed by specifications by subgroup with capture differential associations.

Tables A5 and A6 show the results for science course taking by gender. As before, the first column shows the main specification, followed by the fully interacted models for each gender. Table A7 and A8 show the results by ethnicity.

Table A5 - Results for science course taking by gender

VARIABLES	(1) Main Specification	(2) Male	(3) Female
Female	-0.245***		
	(0.044)		
Black	0.045	0.141	-0.035
	(0.068)	(0.098)	(0.097)
Hispanic	-0.001	-0.009	-0.007
	(0.076)	(0.107)	(0.106)
Asian	0.276**	0.447**	0.115
	(0.105)	(0.145)	(0.146)
multi-race	-0.182	0.019	-0.450*
	(0.109)	(0.126)	(0.180)
Non-science GPA	0.624***	0.565**	0.722***
	(0.133)	(0.176)	(0.203)
Science GPA	-0.163	-0.265	-0.082
(Freshman & Sophomore Years)	(0.134)	(0.181)	(0.202)
EXPLORE Reading	0.015	0.029*	0.001
	(0.009)	(0.012)	(0.013)
EXPLORE Science	0.031**	0.036*	0.024
	(0.012)	(0.016)	(0.017)
EXPLORE Math	0.052***	0.084***	0.021
	(0.009)	(0.012)	(0.013)
EXPLORE Reading	0.021*	0.005	0.035**
	(0.009)	(0.012)	(0.013)
Free/reduced price lunch	-0.187**	-0.203*	-0.154
	(0.063)	(0.090)	(0.090)
Constant	-0.922***	-1.017***	-1.123***
	(0.175)	(0.239)	(0.255)
Observations	3,290	1,622	1,668
R-squared	0.320	0.337	0.318
Cohort FE	YES	YES	YES

Robust standard errors in parentheses

DV = total number of science courses taken about the requirement

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

Table A6 - Results for AP science course taking by gender

VARIABLES	(1) Main Specification	(2) within boys	(3) within girls
Female	-0.046***		
	(0.014)		
Black	-0.061**	-0.025	-0.095**
	(0.022)	(0.030)	(0.033)
Hispanic	-0.069**	-0.029	-0.116**
	(0.025)	(0.034)	(0.036)
Asian	0.027	0.035	0.013
	(0.042)	(0.066)	(0.055)
multi-race	-0.013	0.059	-0.108*
	(0.035)	(0.046)	(0.053)
Cumulative non-Science GPA	0.189***	0.172***	0.196***
	(0.035)	(0.051)	(0.049)
Science GPA	0.040	0.067	0.022
(Freshman/Sophomore Years)	(0.033)	(0.049)	(0.045)
EXPLORE Reading	0.007*	0.006	0.008
	(0.003)	(0.004)	(0.004)
EXPLORE Science	0.020***	0.023***	0.016**
	(0.004)	(0.005)	(0.005)
EXPLORE Math	0.003	0.006	0.000
	(0.003)	(0.004)	(0.004)
EXPLORE English	0.003	-0.001	0.005
	(0.003)	(0.004)	(0.004)
Free/reduced price lunch	-0.013	-0.011	-0.011
	(0.019)	(0.027)	(0.028)
Constant	-0.968***	-1.070***	-0.904***
	(0.052)	(0.069)	(0.079)
Observations	3,290	1,622	1,668
R-squared	0.376	0.388	0.374
Cohort FE	YES	YES	YES

Robust standard errors in parentheses

DV = Indicator for at least one science AP course

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

Table A7 - Results for science course taking by ethnicity

VARIABLES	(1) Main Specification	(2) Within White	(3) Within Black	(4) Within Hispanic	(5) Within Asian	(6) Within multi-racial
Female	-0.245***	-0.146*	-0.338***	-0.179	-0.427*	-0.845***
	(0.044)	(0.061)	(0.090)	(0.111)	(0.190)	(0.235)
Black	0.045					
	(0.068)					
Hispanic	-0.001					
	(0.076)					
Asian	0.276**					
	(0.105)					
multi-race	-0.182					
	(0.109)					
Non-science GPA	0.624***	0.369	0.652**	0.526	1.800*	1.394*
	(0.133)	(0.220)	(0.227)	(0.280)	(0.715)	(0.636)
Science GPA	-0.163	0.106	-0.290	-0.096	-0.734	-0.578
(Freshman & Sophomore Years)	(0.134)	(0.233)	(0.219)	(0.272)	(0.793)	(0.631)
EXPLORE Reading	0.015	0.019	0.017	0.051*	-0.043	0.036
	(0.009)	(0.011)	(0.020)	(0.025)	(0.037)	(0.038)
EXPLORE Science	0.031**	0.020	0.068**	0.027	0.061	-0.064
	(0.012)	(0.016)	(0.025)	(0.031)	(0.070)	(0.057)
EXPLORE Math	0.052***	0.023	0.096***	0.073**	-0.004	-0.002
	(0.009)	(0.012)	(0.020)	(0.023)	(0.040)	(0.044)
EXPLORE Reading	0.021*	-0.003	0.047**	0.041	0.026	0.041
	(0.009)	(0.012)	(0.018)	(0.024)	(0.031)	(0.038)
FRPL	-0.187**	-0.353**	-0.045	0.054	0.282	-0.240
	(0.063)	(0.130)	(0.101)	(0.148)	(0.221)	(0.289)
Constant	-0.922***	0.198	-2.288***	-2.132***	-1.595	0.109
	(0.175)	(0.253)	(0.322)	(0.420)	(0.956)	(0.776)
Observations	3,290	1,697	867	469	119	138
R-squared	0.320	0.128	0.346	0.369	0.369	0.386
Cohort FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses; FRPL = free/reduced price lunch

DV = number of science course above required

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

Table A8 - Results for AP science course taking by ethnicity

VARIABLES	(1) Main Specification	(2) Within White	(3) Within Black	(4) Within Hispanic	(5) Within Asian	(6) Within multi-racial
Female	-0.046***	-0.041	-0.035	-0.067*	-0.008	-0.192**
	(0.014)	(0.022)	(0.019)	(0.029)	(0.088)	(0.073)
Black	-0.061**					
	(0.022)					
Hispanic	-0.069**					
	(0.025)					
Asian	0.027					
	(0.042)					
multi-race	-0.013					
	(0.035)					
Non-Science GPA	0.189***	0.206**	0.130**	0.139*	0.162	0.019
	(0.035)	(0.072)	(0.042)	(0.067)	(0.314)	(0.196)
Science GPA	0.040	0.125	0.024	0.013	0.092	0.264
(Freshman/Sophomore Years)	(0.033)	(0.074)	(0.036)	(0.056)	(0.341)	(0.198)
EXPLORE Reading	0.007*	0.005	0.003	0.018*	-0.030	0.008
	(0.003)	(0.004)	(0.005)	(0.008)	(0.017)	(0.013)
EXPLORE Science	0.020***	0.021***	0.014**	0.013	0.021	0.037*
	(0.004)	(0.006)	(0.005)	(0.009)	(0.023)	(0.018)
EXPLORE Math	0.003	0.004	0.002	0.008	0.019	-0.030**
	(0.003)	(0.005)	(0.004)	(0.007)	(0.015)	(0.011)
EXPLORE English	0.003	0.000	0.005	-0.004	0.028	0.009
	(0.003)	(0.004)	(0.004)	(0.007)	(0.016)	(0.013)
FRPL	-0.013	0.010	-0.039	0.001	-0.110	-0.153
	(0.019)	(0.042)	(0.026)	(0.045)	(0.105)	(0.081)
Constant	-0.968***	-1.328***	-0.649***	-0.836***	-1.014*	-0.862***
	(0.052)	(0.074)	(0.075)	(0.116)	(0.388)	(0.233)
Observations	3,290	1,697	867	469	119	138
R-squared	0.376	0.249	0.259	0.272	0.321	0.491
Cohort FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses; FRPL = free/reduced price lunch

DV = indicator for at least one science AP course

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

Planning

To examine aspirations related to science achievement, we regress whether or not the student professes an intended college major in science on factors of gender, ethnicity and course-taking.

Intended majors were determined by free response items in the senior survey and were coded as 0 for majors not in a traditional science area or 1 for science majors (e.g., chemistry, physics, biology, environmental science). Majors were coded separately if they fell into applied science paths like engineering and medicine. Engineering majors (i.e., any entry with engineering in the name) were coded as 2 and medical ones (e.g., pre-med, veterinary or animal science) were coded as 3. (Nursing and radiology were not counted in this category). However, if students had listed a traditional science major along with an applied (e.g., "Pre-med, biology") the intended major was coded as 1. For the analysis, we choose an inclusive definition of a science major. We collapsed the three codes above into a single dummy variable, where each was counted as a science major. Every other major was coded as 0.

Again, our main specification follows the same form as (1):

$$(4)\ S_{ic} = \alpha + \beta_1 E_{ic} + \beta_2 F_{ic} + \mathbf{X}_{ic}\beta + \delta_c + \varepsilon$$

The term S_{ic} is a dichotomous variable indicating if student *i* intends to declare a science-related major. The vector of student characteristics is the same as in (1).

Table A9 presents the results for planning for gender. As before, the first column present the results from the main specification, followed by the within subgroup results. Table A10 presents the results for ethnicity.

Table A9 - Results for intended science major taking by gender

VARIABLES	(1) Main Specification	(2) Male	(3) Female
Female	-0.029		
	(0.015)		
Black	0.017	-0.025	0.061
	(0.023)	(0.034)	(0.031)
Hispanic	0.000	-0.011	0.011
	(0.026)	(0.038)	(0.034)
Asian	0.032	-0.097	0.139*
	(0.042)	(0.061)	(0.055)
Multi-race	0.107**	0.119*	0.098
	(0.041)	(0.058)	(0.057)
EXPLORE Reading	0.004	0.005	0.003
	(0.003)	(0.004)	(0.004)
EXPLORE Science	0.003	0.002	0.005
	(0.004)	(0.006)	(0.005)
EXPLORE Math	0.004	0.004	0.006
	(0.003)	(0.005)	(0.004)
EXPLORE Reading	-0.005	-0.006	-0.005
	(0.003)	(0.004)	(0.004)
Free/reduced price lunch	0.036	0.057	0.023
	(0.021)	(0.031)	(0.029)
Non-science GPA	-0.135***	-0.127***	-0.156***
	(0.026)	(0.038)	(0.036)
Science GPA	0.126***	0.129***	0.130***
	(0.018)	(0.027)	(0.024)
Total no. science courses	0.046***	0.029***	0.063***
	(0.005)	(0.008)	(0.007)
Constant	-0.132*	-0.065	-0.241**
	(0.063)	(0.093)	(0.084)
Observations	3,333	1,650	1,683
R-squared	0.065	0.049	0.094
Cohort FE	YES	YES	YES

Robust standard errors in parentheses

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

Table A10 - Results for intended science major taking by ethnicity

VARIABLES	(1) Main Specification	(2) Within White	(3) Within Black	(4) Within Hispanic	(5) Within Asian	(6) Within multi-racial
Female	-0.029	-0.040	0.008	-0.085*	0.142	-0.120
	(0.015)	(0.021)	(0.027)	(0.037)	(0.092)	(0.088)
Black	0.017					
	(0.023)					
Hispanic	0.000					
	(0.026)					
Asian	0.032					
	(0.042)					
Multi-race	0.107**					
	(0.041)					
EXPLORE Reading	0.004	0.001	0.001	0.007	-0.027	0.045**
	(0.003)	(0.004)	(0.006)	(0.008)	(0.014)	(0.016)
EXPLORE Science	0.003	0.009	-0.007	-0.008	0.014	0.032
	(0.004)	(0.005)	(0.007)	(0.009)	(0.024)	(0.022)
EXPLORE Math	0.004	0.005	0.005	0.016	-0.001	-0.025
	(0.003)	(0.004)	(0.006)	(0.008)	(0.019)	(0.015)
EXPLORE Reading	-0.005	-0.003	-0.004	-0.012	0.003	-0.026
	(0.003)	(0.004)	(0.005)	(0.007)	(0.015)	(0.014)
FRPL	0.036	0.126**	-0.018	0.063	-0.114	0.009
	(0.021)	(0.044)	(0.031)	(0.046)	(0.089)	(0.092)
Non-science GPA	-0.135***	-0.202***	-0.157***	-0.002	-0.175	0.116
	(0.026)	(0.048)	(0.040)	(0.057)	(0.182)	(0.144)
Science GPA	0.126***	0.182***	0.127***	0.045	0.242	-0.143
	(0.018)	(0.035)	(0.026)	(0.038)	(0.146)	(0.096)
Total science courses	0.046***	0.058***	0.039***	0.039**	0.085*	0.043
	(0.005)	(0.008)	(0.010)	(0.014)	(0.040)	(0.029)
Constant	-0.132*	-0.302**	0.211*	-0.177	-0.558	-0.405
	(0.063)	(0.107)	(0.099)	(0.131)	(0.386)	(0.269)
Observations	3,333	1,717	886	471	120	139
R-squared	0.065	0.081	0.055	0.077	0.240	0.157
Cohort FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses

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

Quasi-experimental design

We use NSHS's science tracking policy (students are tracked into honors, regular, or remedial freshman biology based on their scaled eighth grade reading EXPLORE scores) to examine a number of questions (see bullets below). Tracking based on test scores creates a discontinuity in the probability of participation in three different types of freshman biology classes that enables us to employ a regression discontinuity design to plausibly estimate causal effects of different science tracks. Due to the very small number of students placed into remedial biology, we focus our analysis on placement into honor Biology compared to regular Biology.

Until recently, NSHS used information from student's eighth grade reading EXPLORE scores to place students into honors biology. The cutoff corresponded to a scaled score of 15. Students with a score of 15 or higher were placed into honors biology; those below 15 were placed in regular biology. Critically, this rule was not deterministic. Some students below the cutoff took honors Biology, while some above the cutoff took regular Biology. Thus, we employ what is termed a "fuzzy" regression discontinuity design. This allows us to estimate the causal impact of participation in freshman Honors biology for students just above the cutoff, using the students just below the cutoff as a comparison group. We answer the following research questions:

- Does being placed into honors Biology freshman year increase students' science ACT scores or science GPA?
- Does placement into honors biology enhanced or diminished participation in other science courses, particularly AP courses?
- Are students placed in honors biology more likely to major in a science-related field?

Since placement is not fully determined by eighth grade scores, we employ fuzzy regression discontinuity design using a two-stage least squares approach (Angrist & Pischke, 2008). This method estimates the impact of treatment (placement in honors Biology) by first estimating the likelihood of placement into the treatment due to the score on the qualifying assessment (called the running or assignment variable). Using the fitted values from that estimate (the first stage), the outcome variable is regressed on the fitted values from the first stage to produce an estimate of the impact on treatment. Since this method identifies the causal impact of treatment only for those just above and just below the cutoff, it is known as a "local average treatment effect" or LATE (Angrist & Pischke, 2008).

There are several approaches to determining how many observation to include in the analysis. Researchers must choose a window around the cutoff. A narrower window provides a stronger causal estimate. Students with just a one point on their EXPLORE scores are likely similar on other, unobserved characteristics. However, narrower windows often provide less precision, since they admit fewer observations. Thus, it is common to extend the window to include more observations. But there is a causality/precision tradeoff. In our case, we decided to use a window of ±3 points away from the cutoff of 15. Scores from 12-14 form the comparison group; scores from 15-17 form the treatment group.

The reduced form of the model is:

$$(5)\ A_{ic} = \alpha + \beta_1 P_{ic} + \beta_2 D_{ic} + \mathbf{X}_{ic}\beta + \varepsilon$$

The model is substantially similar to (1) with a few additions. The term P_{ic} is student *i*'s eighth grade reading EXPLORE score, the running variable. The term D_{ic} is an indicator for whether or not a student *i* in cohort *c* had a reading EXPLORE score 15 or greater. This serves as an instrument for placement into honors Biology. The first stage estimates the probability of assignment to the honors level biology class conditional on eighth grade reading scores (the running variable) and a vector of student characteristics (race, gender, income, and prior achievement):

$$O_{ic} = \alpha + \beta_1 P_{ic} + \beta_2 D_{ic} + \mathbf{X}_{ic}\beta + \varepsilon$$

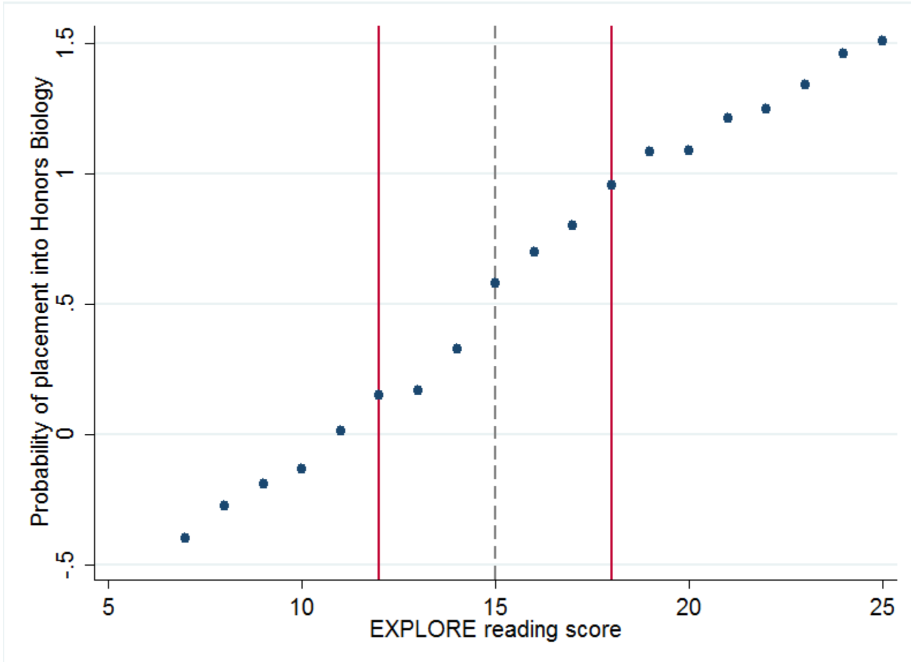
The term O_{ic} is placement into honors Biology. The second stage regresses the achievement outcome on the fitted values from the first stage:

$$A_{ic} = \alpha + \beta_1 \hat{O}_{ic} + \beta_2 P_{ic} + \mathbf{X}_{ic}\beta + \varepsilon$$

The term $\hat{O}$ is the fitted values of O from the first stage regression. The fuzzy RD design allows us to estimate local average treatment effects for students who are just above and just below the assignment cutoff, conditioning on their predicted probability of assignment and actual assignment status.

As noted, the running variable is not a perfect predictor of assignment to treatment at the cutoff. Figure A1 shows the probability of assignment to freshman honors biology based upon middle school EXPLORE reading scores. A score of 15 is the cutoff for taking honors. The vertical red lines represent the lower and upper bound of the window for the instrumental variables regression. Only observations that fell within the window were used. The discontinuity is quite fuzzy, but the probability of assignment jumps at 15 to above 0.5.

The results of the instrumental variables regressions are shown in Tables A11 through A15. In each case, the first column shows the results of the first stage regression, while the second column shows the results of the second stage. The first variable is the variable of interest, assignment to freshman honors biology. The second variable is the instrument for assignment to freshman biology. It is simply an indicator for being above



the cutoff.

Figure A1 - Probability of placement into honors biology

Table A11 - Two-stage least squares results for science ACT

	(1)	(2)
VARIABLES	First Stage	Second Stage
Freshman honors bio		-0.893
		(2.809)
Above cutoff	0.104***	
	(0.036)	
Free/reduced price lunch	-0.127***	-0.943**
	(0.025)	(0.409)
Female	-0.008	-0.896***
	(0.018)	(0.168)
EXPLORE Reading	0.032***	0.236
	(0.009)	(0.163)
EXPLORE Math	0.037***	0.396***
	(0.004)	(0.108)
EXPLORE Science	0.047***	0.484***
	(0.005)	(0.143)
EXPLORE English	0.024***	0.203***
	(0.003)	(0.073)
Black	-0.015	-1.708***
	(0.027)	(0.236)
Hispanic	-0.051	-1.338***
	(0.031)	(0.307)
Asian	0.037	0.277
	(0.046)	(0.428)
Multi-race	0.024	-0.846**
	(0.043)	(0.391)
Constant	-1.395***	5.292
	(0.096)	(3.865)
Observations	1,496	1,496
R-squared		0.525
Cohort FE	YES	YES

Robust standard errors in parentheses

First stage DV = Freshman honors biology; second stage DV = Science ACT score

*** p<0.01, ** p<0.05, * p<0.1

Table A12 - Two-stage least squares results for science GPA, weighted and unweighted

VARIABLES	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage
Freshman honors bio		1.266**		0.755
		(0.641)		(0.563)
Above cutoff	0.104***		0.104***	
	(0.036)		(0.036)	
Free/reduced price lunch	-0.127***	-0.210**	-0.127***	-0.206**
	(0.025)	(0.096)	(0.025)	(0.084)
Female	-0.008	0.191***	-0.008	0.176***
	(0.018)	(0.036)	(0.018)	(0.031)
EXPLORE Reading	0.032***	-0.044	0.032***	-0.033
	(0.009)	(0.036)	(0.009)	(0.032)
EXPLORE Math	0.037***	0.019	0.037***	0.018
	(0.004)	(0.025)	(0.004)	(0.022)
EXPLORE Science	0.047***	0.014	0.047***	0.013
	(0.005)	(0.032)	(0.005)	(0.028)
EXPLORE English	0.024***	-0.002	0.024***	-0.000
	(0.003)	(0.016)	(0.003)	(0.014)
Black	-0.015	-0.431***	-0.015	-0.372***
	(0.027)	(0.054)	(0.027)	(0.047)
Hispanic	-0.051	-0.252***	-0.051	-0.207***
	(0.031)	(0.069)	(0.031)	(0.061)
Asian	0.037	0.200**	0.037	0.178**
	(0.046)	(0.092)	(0.046)	(0.083)
multi-race	0.024	-0.224***	0.024	-0.206***
	(0.043)	(0.076)	(0.043)	(0.067)
Constant	-1.395***	1.962**	-1.395***	1.927**
	(0.096)	(0.876)	(0.096)	(0.767)
Observations	1,496	1,496	1,496	1,496
R-squared		0.494		0.389
Cohort FE	YES	YES	YES	YES

Robust standard errors in parentheses

First stage DV = Freshman honors biology; second stage DV = Science GPA score

Note: GPA is weighted in column 2, unweighted in column 4

*** p<0.01, ** p<0.05, * p<0.1

Table A13 - Two-stage least squares results for science course taking

	(1)	(2)
VARIABLES	First Stage	Second Stage
Freshman honors bio		-0.378
		(1.062)
Above cutoff	0.104***	
	(0.036)	
Free/reduced price lunch	-0.127***	-0.269*
	(0.025)	(0.157)
Female	-0.008	-0.231***
	(0.018)	(0.061)
EXPLORE Reading	0.032***	0.080
	(0.009)	(0.060)
EXPLORE Math	0.037***	0.048
	(0.004)	(0.042)
EXPLORE Science	0.047***	0.066
	(0.005)	(0.055)
EXPLORE English	0.024***	0.004
	(0.003)	(0.028)
Black	-0.015	-0.066
	(0.027)	(0.089)
Hispanic	-0.051	-0.161
	(0.031)	(0.114)
Asian	0.037	0.280*
	(0.046)	(0.158)
multi-race	0.024	-0.170
	(0.043)	(0.167)
Constant	-1.395***	1.198
	(0.096)	(1.491)
Observations	1,496	1,496
R-squared		0.081
Cohort FE	YES	YES

Robust standard errors in parentheses

First stage DV = Freshman honors biology; second stage DV = Number of science course above required

*** p<0.01, ** p<0.05, * p<0.1

Table A14 - Two-stage least squares results for science AP course taking

	(1)	(2)
VARIABLES	First Stage	Second Stage
Freshman honors bio		0.523
		(0.406)
Above cutoff	0.104***	
	(0.036)	
Free/reduced price lunch	-0.127***	-0.012
	(0.025)	(0.060)
Female	-0.008	0.035
	(0.018)	(0.023)
EXPLORE Reading	0.032***	-0.010
	(0.009)	(0.024)
EXPLORE Math	0.037***	0.002
	(0.004)	(0.016)
EXPLORE Science	0.047***	0.003
	(0.005)	(0.020)
EXPLORE English	0.024***	-0.012
	(0.003)	(0.011)
Black	-0.015	-0.130***
	(0.027)	(0.033)
Hispanic	-0.051	-0.110**
	(0.031)	(0.043)
Asian	0.037	0.077
	(0.046)	(0.065)
Multi-race	0.024	-0.062
	(0.043)	(0.059)
Constant	-1.395***	0.142
	(0.096)	(0.558)
Observations	1,496	1,496
R-squared		0.155
Cohort FE	YES	YES

Robust standard errors in parentheses

First stage DV = Freshman honors biology; second stage DV = Indicator for at least one science AP course

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A15 - Two-stage least squares results for science major intention

	(1)	(2)
VARIABLES	First Stage	Second Stage
Freshman honors bio		-0.231
		(0.363)
Above cutoff	0.104***	
	(0.036)	
Free/reduced price lunch	-0.127***	-0.043
	(0.025)	(0.057)
Female	-0.008	-0.023
	(0.018)	(0.021)
EXPLORE Reading	0.032***	0.021
	(0.009)	(0.021)
EXPLORE Math	0.037***	0.020
	(0.004)	(0.014)
EXPLORE Science	0.047***	0.017
	(0.005)	(0.018)
EXPLORE English	0.024***	0.002
	(0.003)	(0.010)
Black	-0.015	0.036
	(0.027)	(0.033)
Hispanic	-0.051	0.050
	(0.031)	(0.041)
Asian	0.037	0.146**
	(0.046)	(0.065)
Multi-race	0.024	0.096
	(0.043)	(0.059)
Constant	-1.395***	-0.378
	(0.096)	(0.504)
Observations	1,496	1,496
R-squared		-0.026
Cohort FE	YES	YES

Robust standard errors in parentheses

First stage DV = Freshman honors biology; second stage DV = Indicator for intending a science related major

*** p<0.01, ** p<0.05, * p<0.1