Low Performance in Math and English: Do Women React Differently than Men?

Paola Ugalde A

Arizona State University

Revision: March 24, 2023

Abstract

Career advancement requires an individual to overcome setbacks at every stage in life. If women and men react to them differently, an early defeat might preclude a woman from advancing in a given area, which could be an explanation for women being under-represented in high-level positions. I study the gender difference in reaction to low performance by focusing on the decision of North Carolina public high school students to enroll in advanced math or English classes after learning about their performance on statewide standardized tests in each subject. I find that women react more strongly than men to low performance in math relative to English. Given the common association of men with math and women with English, the results suggest that the gender stereotype of the area in which a woman faces a difficulty might be relevant for her reaction.

1 Introduction

Women are underrepresented in high-level positions in many areas. They represented 31% of the senior management positions in 2021 (Grant Thornton International Ltd, 2021) and held only 7.4% of the CEO positions in the Fortune 500 companies in 2020 (Pew Research Center, 2021). Making it to the top requires people to overcome setbacks at every stage in life, from low grades and college admission rejections, to bad job interviews or performance evaluations. If women react more adversely to these setbacks than men, an early defeat might preclude a woman from advancing.

In this paper, I study the gender difference in reaction to low performance. I focus on the decision of high school students to enroll in advanced math or English classes after learning about their performance on subject-specific standardized tests. The idea is to determine whether women react significantly more to low performance on those tests than men, and if the reaction varies with the subject. Specifically, I establish if low performance differentially impacts the probability that a woman or a man enrolls in advanced classes.

There is a growing literature that concludes that women are more likely to drop out of male-dominated majors like Economics and STEM (Science, Technology, Engineering, Mathematics) after experiencing a difficulty early in their college careers. In Goldin (2015), the author documents that women getting B- or less in their introductory economics course are less likely to graduate with an economics major than men with the same final grades. Similarly, Rask and Tiefenthaler (2008) estimate that the probability of a woman enrolling in a second course in economics decreases if she gets grades in the lowest quartile of the distribution during her first economics class, but the same is not true for men. In Kugler et al. (2021), the authors conclude that if women believe that men are inherently a better fit for STEM majors, and if there are less women than men in their major, they are more likely to perceive their low grades as confirmation of their unfitness for their male-dominated STEM major and drop out. Further, Ahn et al. (2019) estimate a structural model in which women care more about grades and conclude that if women had the same preferences for grades as men the gender gap in STEM majors would close by 8%. However, a question that remains open is whether the reaction to setbacks is the same in male and female-dominated fields. Furthermore, given that most of the existing evidence is at the higher education level, analyzing the reaction to setbacks earlier in life can provide insight about critical decisions that occur later on.

The interest in the difference between subjects lies in the common association of men with math and women with English and languages. For example, according to the National Center for Education Statistics (2018), in 2015 more male than female students identified math as their favorite subject, and more female than male students reported reading as their favorite activity. Riegle-Crumb and Humphries (2012) concludes that teachers believe math is easier for white boys than for white girls. Women are less confident in their math abilities (Ellis et al., 2016, Ganley and Lubienski, 2016) and have a larger advantage over men in verbal skills (Aucejo and James, 2021; Delaney and Devereux, 2019; Breda and Napp, 2019). However, taking more high school math courses increases wages for female college graduates and the likelihood of a woman majoring in "more technical and nontraditional fields" (Levine and Zimmerman, 1995)[1].

I study the course-taking behavior of North Carolina public high school students for the cohorts starting 9^{th} grade in the 2013 and 2014 scholar years. The North Carolina Department of Public Instruction (NCDPI) uses standardized End-of-Course (EOC) tests to assess students' knowledge about specific subjects for high school accountability purposes (North Carolina Department of Public Instruction, 2019b). During high school, every student must take an EOC test at the end of the first math course (Math 1), the second English course (English 2) and the Biology course. For each subject tested, there is a threshold score above which students are considered proficient, i.e. they have sufficient command of skills for the respective course and are prepared for further studies in the subject (North Carolina Department of Public Instruction, 2016a). I focus on the EOC tests for Math 1 and English

[1] In Levine and Zimmerman (1995), a nontraditional female field is a major in which more than 70% of the students are men.

2, and the corresponding choice of regular versus advanced class in Math 2 and English 3.

In North Carolina, each spring semester students make their choices for the classes they will be taking the next year. Often, students decide between the regular or honors version of a class, and when available they can consider taking Advanced Placement (AP) and International Baccalaureate (IB) classes. These advanced classes are more difficult (AP and IB are college-level classes), and cover more topics than the regular class. The benefits associated with taking advanced classes during high school are numerous. For example, taking advanced courses is related to increases in the probability of high school graduation, 4-year college attendance and college graduation, and with the completion of more college credits and a higher GPA during college (Long et al., 2012). Moreover, taking advanced math classes has a positive effect on earnings later in life (Rose and Betts, 2004; Joensen and Nielsen, 2009). Therefore, it is important to understand if the advanced course-taking behavior of high school students of different genders is influenced differently by the perception of low performance on a test.

The existence of the proficiency cutoffs on the North Carolina EOC tests suggests the implementation of a regression discontinuity (RD) design. However, the main interest is not in the discontinuity generated by the cutoff that quantifies the effect of low performance on the likelihood of enrolling in an advanced class, but in the gender difference in the discontinuity. Thus, I use a difference-in-discontinuity approach that combines regression discontinuity (RD) and differences-in-differences (DD) in order to identify the gender differences in the discontinuity created by the cutoffs.

I find that the effect of low performance is not the same for math and English. Women react more to low performance in math than in English relative to men. The probability of enrolling in honors Math 2 changes with performance for women, but not for men. Women that perform poorly on the Math 1 EOC tests are 5 to 7 percentage points less likely to enroll in honors Math 2 while men's likelihood is the same regardless of performance status. On the other hand, the effect of low-performance on the English test is smaller than in math. As in the math case, men do not seem to react to low-performance. However, women's probability of enrolling in Advanced English 3 is between 2 and 4 percentage points lower among non-proficient students. The fact that women react more to low performance in math than in English suggests that the area in which a woman is facing a difficulty might be relevant for her reaction. This result is in line with Kaganovich et al. (2021), which concludes that women's grade sensitivity depends on the area or category.

My primary results focus on advanced course-taking decisions for Math 2 and English 3, however I also analyze the effect of low performance on higher-level advanced course-taking behavior at the end of high school. I find that the number of higher-level advanced courses taken by non-proficient women is smaller relative to non-proficient men. Roughly, these differences translate to a reduction of 6% and 14% in the number of higher-level advanced courses taken by women in English and math, respectively. These results are in line with Tan (2020) which finds that college students at the margin of two letter grades that receive the worse grade are more likely to take easier classes in the subsequent years.

There is evidence that test scores can influence students' decisions about post-secondary education (Papay et al., 2016). Therefore, I study the effect of low performance on the plans to attend a 4-year college.[2] I find that women are 2 and 3 percentage points less likely to make plans of going to college than non-proficient men when deemed not proficient on the English 2 and Math 1 EOC tests, respectively.

My findings are consistent with women interpreting their non-proficiency status as a discouraging signal of their ability, and this may affect their future behavior in terms of class enrollment and college attendance. There is evidence that women attribute negative feedback to lack of ability, update their beliefs more pessimistically, and are less willing to compete than men (Roberts and Nolem-Hoeksema, 1989; Shastry et al., 2020; Berlin and Dargnies, 2016; Buser and Yuan, 2019). All of these traits could lead to women avoiding "harder" classes after performing poorly on the standardized tests. Additionally, the fact that

[2] I only observe what students plan to do after high school, but I do not know if they executed those plans.

the effect of low performance is stronger in math than in English highlights the importance of the stereotype associated with a given domain, which is in line with results from laboratory settings in which women and men respond to feedback differently depending on the gender stereotype associated with a task (Kiefer and Shih, 2006; Coffman et al., 2019).

The rest of the paper is organized as follows. Section 2 discusses the North Carolina standardized testing system and the administrative data used in the analysis. Section 3 explains the empirical strategy. Sections 4 and 5 present the results. Section 6 discusses possible mechanisms that could be driving the results. Finally, Section 7 concludes.

2 Institutional Background and Data

The North Carolina Department of Public Instruction (NCDPI) uses standardized end-of-course (EOC) tests to "sample a student's knowledge of subject-related concepts as specified in the North Carolina Standard Course of Study and to provide a global estimate of the student's mastery of the material in a particular content area" (North Carolina Department of Public Instruction, 2019b) for high school accountability purposes. During high school, every student must take an EOC test at the end of the first math course (Math 1), the second English course (English 2) and the Biology course. I focus on Math 1 and English 2 tests, usually taken during 9^{th} and 10^{th} grade, respectively. Each of these tests represent 20% of the final course grade (North Carolina Department of Public Instruction, 2016b, 2016c).

Based on the test scores, students are classified into five achievement levels, where one is the lowest achievement level and five the highest. Students that receive at least an achievement level of 3 are considered proficient, which means that they have sufficient command of skills for the respective course and are prepared for further studies in that subject (North Carolina Department of Public Instruction, 2016a). For each subject, there is an established threshold above which students are considered proficient. These clear cutoffs allow for the application of the empirical strategy described in Section 3.

Most high schools in North Carolina employ a block schedule or semester plan, in which students take four classes each semester, for a total of eight per year (Averett, 1994). High school graduation requirements in North Carolina include the completion of four math credits and four English credits, which are equivalent to four courses in each subject, usually taken one each year of high school.[3] The block schedule and adequate progress from grade to grade make it possible to study each subject only one semester each high school year, for instance, by taking English during fall semesters and math during spring semesters.[4]

Each scholar year, during the spring semester (around March or April), students make the choices for the classes they will be taking the next year. When choosing the classes, students often must decide whether to enroll in a regular or honors version of a given course. This is the case for Math 2, for instance. The honors version is a class with a higher level of difficulty, which studies the topics in a deeper way than the regular version and sometimes covers more topics. Given the higher level of difficulty, students taking honors classes get quality points that make grades from honors classes have a higher weight in GPA calculations. For example, a C in a honors class is equivalent to a B in a regular class (North Carolina Department of Public Instruction, 2022).

Once students advance to higher grades, the available options can be more than just honors and regular. For example, the NCDPI allows for Advanced Placement (AP) classes in English Language and Literature or, if offered, an International Baccalaureate (IB) English class to count as the third and fourth English credits. AP and IB courses are college-level classes in which students can earn college credit depending on performance on a test at the end of the course. High school students get even more quality points for taking AP and IB courses than for honor classes.

However, quality points and more knowledge are not the only benefits of taking advanced classes. For instance, in Long et al. (2012), the authors conclude that taking more rigorous

[3]Students can take more than the four required courses in each subject.

[4]However, this is not the only possibility, the block schedule and the option to take high school level classes during middle school allow for different paths. For example, four math credits can be obtained by taking (and passing) one math class each semester during the first two years of high school.

courses, in math, English or science, during high school increases the probability of high school graduation and 4-year college attendance.[5] Additionally, college students who took advanced classes during high school tend to complete more college credits, have a higher GPA, and a higher probability of graduation. Moreover, studies like Rose and Betts (2004) and Joensen and Nielsen (2009) find that taking advanced math classes during high school has a positive effect on earnings later in life.[6]

In order to determine if there is a gender difference in the reaction to low performance on EOC tests, I use administrative records from the North Carolina Education Research Data Center (NCERDC). My analysis focuses on the cohorts that began public high school on the fall of 2013 and 2014. Since one of my main objectives is to study the effect of test scores on class choices for the next scholar year, I restrict the samples to those students taking the relevant EOC tests during the fall semesters. In this way I guarantee that students have received the exam results prior to making course decisions. Additionally, I restrict the samples to students for which the relevant transcript information for the following school year is observed.

The Math 1 sample includes 27,997 students that began high school in the 2013-2014 or 2014-2015 school years and took the Math 1 EOC test at the end of the fall semester during their freshman year.[7] The English 2 sample includes 72,395 students that began 10th grade in the fall of 2013 or 2014 and that took the English 2 EOC test at the end of that semester.[8] Table 1 shows summary statistics for both samples. They are balanced in terms of gender with 51% of women. More than half the students are white (54% math, 57% English), around

[5]The National Center for Education Statistics (2019) report about math, science and reading instruction finds a positive correlation between taking advanced math classes and 4-year college acceptances.

[6]In Joensen and Nielsen (2009), the increase in earnings is due to a higher probability of attending college.

[7]About 16% of the students took the Math 1 class during the fall of their freshman year, 51% took it after the fall semester and 33% took it during 8th grade (some middle schools offer 9^{th} grade-level math classes, which allows students to gain high school credits and take 10^{th} classes during their freshman year). Fall students performed better in academic terms than the after fall, but worse than the students that took the class during middle school. See Table A1.

[8]Around 38% of the students took the English 2 class during the fall of their sophomore year, 62% took it after the fall semester and 7% during their freshman year. There are not economically significant differences between the fall and after fall students. See Table A2

Table 1: Sample Characteristics

	Math Sample	English Sample
Female	0.51	0.51
Black	0.26	0.24
Hispanic	0.13	0.11
White	0.54	0.58
EDS	0.51	0.42
Non-Prof. Middle School Math test	0.63	0.35
Non-Prof. Middle School Reading test	0.53	0.38
Non-Prof. Math 1 test	0.49	-
Non-Prof. English 2 test	-	0.35
N	27,997	72,395

Note: Table presents sample proportions of variables of interest. Math sample includes students that took the Math 1 class during the fall semester of their freshman year. English sample includes students that took the English 2 class during the fall semester of their sophomore year. Non-Prof.: non-proficiency, EDS: Economically Disadvantaged Student. *Significant at 10%, **5%, ***1%.

25% Hispanic and 12% or less are black. The proportion of economically disadvantaged students (EDS) is 52% and 43% for the math and English samples, respectively. In terms of academic outcomes, 63% (35%) of the students in the math (English) sample were deemed to proficient on their middle school math test. On the middle school reading test these numbers are 53% and 38% for the math and English samples, respectively.[9] Finally, there is a sizable proportion of students that performed poorly on the tests of interest in the respective samples: 49% of the math sample was deemed not proficient on the Math 1 test, and 35% of the students in the English sample were not proficient on the English 2 test.

Additionally, I am interested in the effect of low performance in outcomes closer to the end of high school like the number of higher-level advanced classes in math and English that students take, and plans to attend a 4-year higher education institution. Therefore, I further restrict the samples to students that graduated at the end of their fourth year for this part

[9]The reading and math tests during 8^{th} are part of the end-of-grade (EOG) exams that North Carolina students take at the end of the year from grades 3 to 8 in order to measure their performance on "the goals, objectives, and grade-level competencies specified in the North Carolina Standard Course of Study" (North Carolina Department of Public Instruction, 2019a).

of the analysis.[10]

3 Empirical Strategy

Given that the main objective is to determine if there is a gender difference in the reaction to low performance on EOC tests, and the existence of a cutoff below which students are considered non-proficient in that subject, I use a difference-in-discontinuity approach that combines regression discontinuity (RD) and differences-in-differences (DD).[11] The RD part exploits the cutoff and requires the identifying assumption that students just above and just below the cutoff are very similar except for the difference in proficiency status. The RD design estimates the size of a discontinuity, in other words it identifies the effect of low performance on an EOC test on the likelihood of enrolling in the advanced version of the class for the next school year. The DD part allows me to determine if there is a gender difference in the discontinuity, i.e. if there is a gender difference in the effect of low performance.

This difference in the discontinuity can be estimated by the interaction of the treatment (being non-proficient) and a gender variable. In practice, I estimate the following model:

$$\begin{aligned} Y_{ijkt} =& \beta_0 + \beta_1 \mathrm{F}_i + \beta_2 \text{Non-Prof}_{ijkt} + \beta_3(\mathrm{F}_i \cdot \text{Non-Prof}_{ijkt}) \\ &+ f(S_{ijkt}) + \text{Non-Prof}_{ijkt} \cdot f(S_{ijkt}) + \gamma \mathbf{X}_{ij} + \eta_j + \nu_t + \epsilon_{ijkt} \end{aligned} \tag{1}$$

where $k \in \{\text{math, English}\}$, and Y_{ijkt} is the outcome variable for student i at high school j that took the test for subject k during school year t. F_i is an indicator variable equal to one for females. Non-Prof_{ijkt} is an indicator that takes value one when the student is not proficient on the EOC test. η_j and ν_t are high school and year fixed effects, respectively. $\mathbf{X}$ includes controls like middle school test scores, race, EDS. $f(S_{ijkt})$ represents a function of the EOC test score, which is the running variable in this case. The interaction between $f(S_{ijkt})$ and

[10] The main results are robust to use the restricted samples.

[11] See Buser and Yuan (2019), Grembi et al. (2016); Eggers et al. (2018); Galindo-Silva et al. (2020) for examples of this design.

Non-Prof$_{ijkt}$ allows for different slopes above and below the cutoff. The main results in the next sections are estimated using $f(S_{ijkt})$ as a first-degree polynomial, however all the results are robust to using a second-degree polynomial instead. The parameter of interest is β_3, the coefficient on the interaction between being a female and non-proficient that estimates the gender difference in the discontinuity. In all regressions, I follow the recommendation of Lee and Card (2008) for regression discontinuity with a discrete assignment variable and cluster the standard errors at the test score level.

3.1 Validity of Design

Given that RD is part of the difference-in-discontinuity design, it is important to make sure that the running variable is exogenous, in other words that test scores are not precisely manipulated. Additionally, all other factors that play a role when deciding whether to take the advanced or regular version of a class are continuously related with the test scores, but more importantly they do not change differentially across genders. These assumptions guarantee that groups of students above and below the proficiency threshold are similar in terms of relevant outcomes and characteristics.

3.1.1 Test Score Manipulation

An exogenous running variable is required because manipulation of the test scores can lead to identification problems. I apply the tests described in McCrary (2008) and Frandsen (2017) where the idea is to test the continuity of the running variable density at the cutoff. For both subjects and both genders, the hypothesis of manipulation cannot be rejected at usual significance levels, regardless of the test.[12] However, it is important to remember that

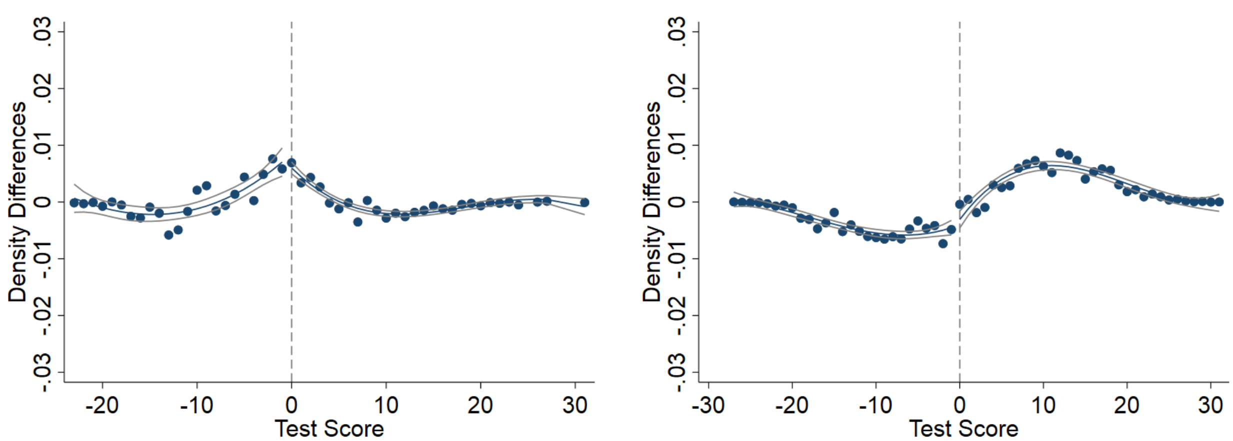
[12]These results are not surprising given the distribution of the test scores in Figure A1. In both cases, there are spikes in the distribution around the cutoff. However the spikes seem to be similar across genders in both subjects, which suggests that if any manipulation exists, it is similar for men and women. Another explanation for the shape of the distributions is the performance measurement system that was in place in North Carolina at the time. Under that system, high school performance depended heavily on the percentage of students considered proficient on the EOC tests (North Carolina Department of Public Instruction, 2015). Such proficiency-count systems create strong incentives for the teachers to direct resources and attention to

the validity of the design is only compromised when the agents can "precisely" manipulate the running variable (Roberts and Whited, 2013), and this kind of manipulation is dubious in the case of the EOC tests in North Carolina.

For instance, one can think that students have some control over their scores but they cannot predict them with certainty or marginally change them around the cutoff. There are several versions of the test, and the exact number of correct answers required to achieve proficiency is not public information and varies by test version. Given that the test scores are used to assess teacher and high school performance, one might be concerned that the teachers are manipulating the results. However, this seems very implausible because tests are not graded by the professor. Instead the NCDPI has in place a centralized grading system with a rigorous protocol to ensure the security of the materials before and during the tests and to avoid any manipulation when transporting and scanning the answer sheets at the NCDPI offices (North Carolina Department of Public Instruction, 2016b, 2016c).[13]

Figure 1: Test Score Densities Differences

(a) Math 1. (b) English 2.



Note: This is a visual test of the continuity of the difference between female and male densities at the cutoff. The central lines are third-degree local polynomials fitted to the data separately above and below the threshold, the lateral lines represent 95% confidence intervals. Scores are normalized such that a score of 0 or more means proficiency.

students on the margin of being above the proficiency threshold, and to pay less attention to the tails of the distribution (Macartney et al., 2018).

[13] Dee et al. (2019) documents the elimination of teacher test score manipulation once a centralized grading system was adopted in New York.

On the other hand, given that the interest is in the gender difference in the discontinuity, it is important to guarantee that if any manipulation exists, it is similar across genders. Following the visual test proposed by Grembi et al. (2016), Figure 1 plots the difference between male and female test score densities for Math 1 and English 2, along with a third-degree polynomial fitted to the data separately above and below the threshold and 95% confidence intervals. These figures support the assumption that the difference in the densities across genders is continuous at the cutoff.

3.1.2 Continuity of Predetermined Covariates

Given that the main interest is the gender difference in the discontinuity, it is important to rule out the possibility of discontinuities in the predetermined covariates varying by gender. In order to do so I estimate model (2) for each subject $k \in \{\text{math, English}\}$ with each of the predetermined controls included in $\mathbf{X}$ as dependent variables, within the optimal bandwidth for each subject used in the main results (± 6 for math and ± 4 for English).

$$\begin{aligned} X_{ijk} = &\beta_0 + \beta_1 \mathrm{F}_i + \beta_2 \text{Non-Prof}_{ijk} + \beta_3 (\mathrm{F}_i \cdot \text{Non-Prof}_{ijk}) \\ &+ f(S_{ijk}) + \text{Non-Prof}_{ijk} \cdot f(S_{ijk}) + \epsilon_{ijk} \end{aligned} \tag{2}$$

Results are presented in Tables 2 and 3 for math and English, respectively. There are no gender differences for any of the math covariates, and the gender difference in the proportion of black students in the English 2 sample is small (2%), however the main results control for all these covariates in order to avoid any biases due to the discontinuities.

The continuity of the covariates, ignoring the possibility of gender differences, is studied by estimating instead model (3) for each subject $k \in \{\text{math, English}\}$ with each of the predetermined controls included in $\mathbf{X}$ as dependent variables.

$$X_{ijk} = \beta_0 + \beta_1 \text{Non-Prof}_{ijk} + f(S_{ijk}) + \text{Non-Prof}_{ijk} \cdot f(S_{ijk}) + \epsilon_{ijk} \tag{3}$$

Table 2: Continuity of the Covariates for Math 1 Test by Gender

	Middle School Math test	Middle School Reading test	Black	Hispanic	EDS
	(1)	(2)	(3)	(4)	(5)
Female (F)	-0.030***	0.150***	0.015	0.009	-0.001
	(0.006)	(0.018)	(0.013)	(0.007)	(0.009)
Non-Prof.	-0.008	-0.010	-0.006	0.018	-0.005
	(0.024)	(0.021)	(0.010)	(0.013)	(0.011)
F*Non-Prof.	-0.017	0.019	0.012	-0.020	0.012
	(0.013)	(0.027)	(0.013)	(0.012)	(0.015)
Mean	-0.10	-0.07	0.24	0.13	0.51
N	14,184	14,152	14,918	14,918	14,918

Note: Dependent variable indicated at the top of each column. Each column follows the same specification: dependent variable regressed on a variable equal to one when the student is a woman, a variable equal to one when the student is deemed non-proficient on the Math 1 EOC test, the interaction of those two and a first degree polynomial of the test score with flexible slopes above and below the proficiency cutoff and high school and year FE. Test scores are standardized such that mean is zero and standard deviation is one. EDS: Economically Disadvantaged Student. Bandwidth of ±6 around the cutoff. Standard errors are clustered at test score level. *Significant at 10%, **5%, ***1%.

Table 3: Continuity of the Covariates for English 2 Test by Gender

	Middle School Math test	Middle School Reading test	Black	Hispanic	EDS
	(1)	(2)	(3)	(4)	(5)
Female (F)	-0.184***	-0.101***	0.047***	0.002	0.045**
	(0.016)	(0.009)	(0.008)	(0.007)	(0.014)
Non-Prof.	-0.025	-0.000	0.001	-0.011	-0.016
	(0.040)	(0.013)	(0.006)	(0.011)	(0.009)
F*Non-Prof.	-0.013	0.002	-0.022**	0.007	0.015
	(0.020)	(0.011)	(0.009)	(0.008)	(0.014)
Mean	-0.15	-0.19	0.28	0.13	0.49
N	19,963	19,966	21,472	21,472	21,472

Note: Dependent variable indicated at the top of each column. Each column follows the same specification: dependent variable regressed on a variable equal one when the student is a woman, a variable equal to one when the student is deemed non-proficient on the English 2 EOC test, the interaction of those two and a first degree polynomial of the test score with flexible slopes above and below the proficiency cutoff, and high school and year FE. Test scores are standardized such that mean is zero and standard deviation is one. EDS: Economically Disadvantaged Student. Bandwidth of ±4 around the cutoff. Standard errors are clustered at test score level. *Significant at 10%, **5%, ***1%.

The results are presented in Figures A2 and A3 in the Appendix for math and English, respectively. They suggest that there are no statistically significant discontinuities at the cutoff except for the proportion of black students in the English sample (in line with the results of the previous analysis).

Overall, I find no evidence of economically significant discontinuities in covariates at the threshold for both math and English tests, which provides some evidence that supports the chosen empirical strategy and the results presented in the next section.

4 Early High School Outcomes

This section presents the main results of the study. Discussion centers on the estimation of model (1) where the outcomes of interest are the decisions to enroll in advanced math or English classes the following year after taking the math or English EOC test, respectively. First, I analyze how the gender difference changes when adding different controls for a specific bandwidth and identify the preferred specification. Then, I examine how the main effect changes (or not) when varying the bandwidths in the preferred specification.

4.1 Math

Table 4 shows the results of the estimation of model (1) for math within a bandwidth of ± 6 points above and below the proficiency cutoff. The optimal bandwidth is between 6 and 7, according to the procedures suggested by Imbens and Kalyanaraman (2012) and Calonico et al. (2019).[14] The outcome variable equals one when the student enrolls in honors Math 2, zero otherwise. The coefficient of interest is the interaction between being a woman and getting a non-proficient score on the Math 1 EOC test (F*Non-Prof). This coefficient estimates the effect of being a non-proficient woman on the likelihood of enrolling in honors Math 2. In other words, it estimates the gender difference in the discontinuity.

[14] The results for a bandwidth of ± 7 can be found in Appendix Table A4.

The differences across specifications in Table 4 are the controls included in each case. The sample size differences arise because it was not possible to recover the middle school test scores for all the students in the math sample. Nonetheless, the proportion of students enrolling in honors Math 2 is around 23% across all specifications.

First, note that the coefficient for being non-proficient is not statistically different from zero and small for most of the specifications. This suggests that the probably of men enrolling in honors Math 2 does not change due to the proficiency level they receive on the Math 1 EOC test. On the other hand, the coefficient for female is positive and significant in all specifications. This means that women are more likely to enroll in honors Math 2 than men. This result, along with the analogous conclusion for English in the next subsection, confirms a pattern already documented in the literature (Shettle et al., 2007; Nord et al., 2011; Long et al., 2012): women are more likely to take advanced courses during high school. The female coefficient can be understood as the gender gap in the probability of enrolling in honors Math 2 for proficient students. Hence, a negative coefficient for being a non-proficient woman implies that although women are more likely to enroll in honors Math 2, the likelihood is not the same for non-proficient women. In other words, this coefficient represents the gender difference in the reaction to low performance and it implies that non-proficient women are less likely to enroll in honors Math 2, but the same is not true for non-proficient men.

A visual representation of this results is shown in Figure 2.[15] The proficiency cutoff is represented by zero. To the left of zero students are considered non-proficient. There is no discontinuity at the cutoff for men (green solid line). However there is a discontinuity at the cutoff for women (maroon dashed line), non-proficient women are less likely enroll in honors Math 2 than proficient ones.

In Table 4, the coefficient on the interaction between being a woman and receiving a non-proficient score on the test varies between -6.5 and -4.9 percentage points depending on the

[15]This is the graphical representation of the specification in column (1) in Table 4.

Table 4: Probability of Taking Honors Math 2. ±6 bandwidth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female (F)	0.119***	0.108***	0.108***	0.108***	0.107***	0.108***	0.108***	0.110***
	(0.014)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)
Non-Prof.	-0.005	-0.006	-0.006	-0.006	-0.006	-0.007	-0.007	-0.008
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)	(0.006)
F*Non-Prof.	-0.065***	-0.054***	-0.053***	-0.052***	-0.052***	-0.049***	-0.049***	-0.050***
	(0.016)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)
Math Test MS						0.101***	0.100***	
						(0.009)	(0.008)	
Reading Test MS							0.003	
							(0.006)	
(Math-Reading) Test MS								0.028***
								(0.005)
School FE		✓	✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
EDS				✓	✓	✓	✓	✓
Race					✓	✓	✓	✓
Mean	0.23	0.23	0.23	0.23	0.23	0.23	0.23	0.23
R^2	0.06	0.22	0.22	0.23	0.23	0.25	0.25	0.23
N	14,918	14,918	14,918	14,918	14,918	14,184	14,146	14,146

Note: The dependent variable is the same across all specifications: one when taking honors Math 2, zero otherwise. Each specification includes indicator variables for being a female (F), for being non-proficient (Non-Prof) on the Math 1 EOC test, and the interaction of those two variables. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Reading Test MS: standardized middle school reading test score. Math test MS: standardized middle school math test score. (Math-Reading) Test MS: difference between the standardized math and reading middle school test scores. School FE: school fixed effects. EDS: Economically Disadvantaged Student. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.

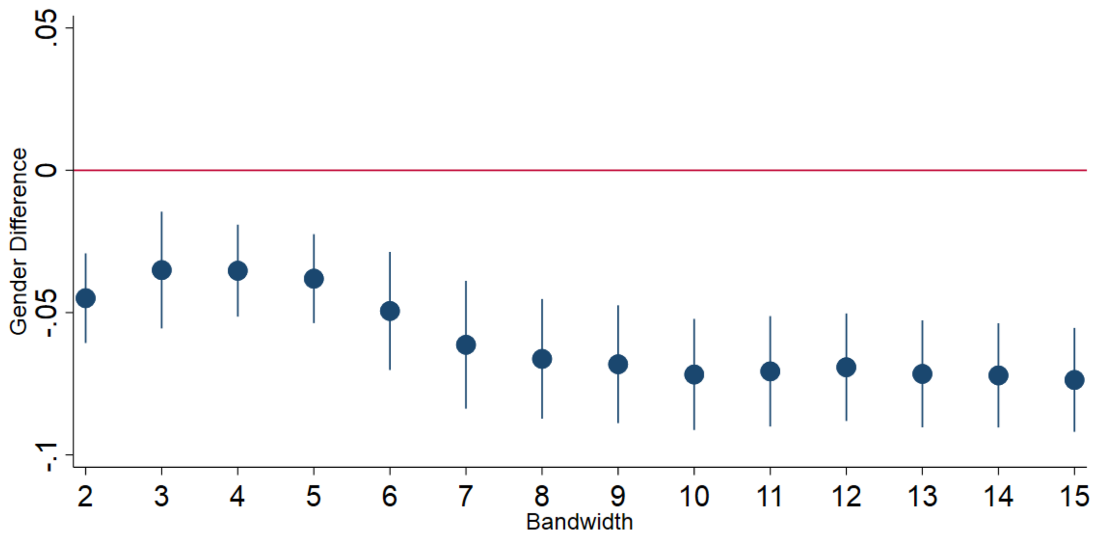
Figure 2: Regression Discontinuity, Math

Note: Test scores are normalized such that a score of 0 or more means proficiency. Lines are fitted values from regressing a dependent variable that equals one when taking honors Math 2 on indicators for being female and non-proficient on the Math 1 EOC test, the interaction of those two variables, and a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, within ±6 of the cutoff. Standard errors are clustered at test score level. The markers represent the proportion of students by gender that take honors Math 2 within each 1 point bin. *Significant at 10%, **5%, ***1%.

specification. High school fixed effects (FE) account for differences across high schools that can affect the likelihood of enrolling in honors classes (for example, high school culture about promoting the advanced classes) and they reduce the gender difference in the discontinuity by about one percentage point. Adding year fixed effects and controlling for race and EDS does not change the effect of being a non-proficient women a considerable amount, which is expected given that these variables do not show a discontinuity at the threshold (See Table 2).

Columns (6)-(8) in Table 4 control for middle school test scores in math and reading. The objective is to control for the ability of the students in these two subjects right before starting high school, and possible comparative advantages. The literature suggests that women have a comparative advantage in verbal skills, while men have comparative advantage in math (Aucejo and James, 2021; Breda and Napp, 2019; Delaney and Devereux, 2019). However, the coefficient of interest does not change in a meaningful way when past test scores are controlled for, each one individually or when the difference between the two is accounted for. Nonetheless, the middle school math test score seems to be relevant in explaining the regular versus honors decision for Math 2, the higher the score the more likely the student is to enroll in the honors class.

Figure 3: Estimates of the Difference in the Discontinuity for Different Bandwidths, Math



Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female in the preferred specification: an indicator for taking honors Math 2 regressed on indicators for female and being non-proficient on the Math 1 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, year and high school FE, indicators for EDS, race and controls for middle school test scores. Standard errors are clustered at test score level. Optimal bandwidth is between 6 and 7.

Results in Table 4 are qualitatively robust to different bandwidths.[16] The preferred specification controls for high school and year FE, EDS, race, and middle school test scores (column (7) in Table 4). Figure 3 shows the estimates of the gender difference in the reaction to non-proficiency for different bandwidths around the cutoff, for this specification.[17] Regardless of the bandwidth chosen, the estimates of the gender difference are consistently between 4 and 7 percentage points, and statistically significant in all cases. This means that women react significantly more strongly to low performance on the Math 1 EOC test than men. Comparing these estimates to the proportion of women that score within 6 points of the cutoff and take honors Math 2 (27%), roughly translates the effect of getting a non-proficient score to a 18 to 24% reduction in the likelihood of enrolling in the honors version of Math 2.

[16]Appendix Tables A3 and A4 show the analogous results for bandwidths ±5 and ±7, respectively.
[17]See Appendix figures A4, and A5 for different specifications.

4.2 English

Table 5 presents the results of the estimation of model (1) for English within a bandwidth of ±5 from the cutoff. The optimal bandwidth is between 4 and 5, according to the procedures suggested by Imbens and Kalyanaraman (2012) and Calonico et al. (2019).[18] The outcome variable is equal to one when the student enrolls in the advanced version of English 3, zero otherwise. Advanced English 3 includes the honors class and AP or IB classes when offered. As in the math case, the coefficient of interest is the interaction between being a woman and getting a non-proficient score on the English 2 EOC test, (F*Non-Prof). This coefficient is an estimate of the effect of being a non-proficient woman on the probability of enrolling in advanced English 3. Hence it represents the estimated gender difference in reaction to low performance.

The difference across the specifications in Table 5 are the controls included in each case, which follows the same pattern as in the math results. The sample size reduces when controlling for final grades and middle school test scores for the same reason as in math, it was not possible to recover them for all the students in the sample. Nonetheless, the proportion of students choosing the advanced version of English 3 is constant across all specifications, around 42%.

Regardless of the specification, the coefficient for being non-proficient is always close to zero and not statistically significant. This suggests that the probability of men enrolling in advanced English 3 does not change because of a non-proficient result on the English 2 test. On the other hand, the coefficient for female is always positive, which suggests women are more likely to enroll in advanced English classes than men. However, the coefficient for the interaction between being a woman and getting a non-proficient score on the English 2 EOC test is not always statistically significant, which suggests that in the case of English women and men might not be reacting differently to a non-proficient score. This result is illustrated

[18] Results for ±4 and ±6 can be found in Appendix Tables A6 and A7, respectively.

Table 5: Probability of Taking Advanced English 3. ±5 bandwidth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female (F)	0.108***	0.097***	0.097***	0.103***	0.104***	0.118***	0.134***	0.111***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Non-Prof.	0.008	0.011	0.011	0.009	0.009	0.009	0.011	0.010
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.007)	(0.009)
F*Non-Prof.	-0.018	-0.022*	-0.022*	-0.020*	-0.020*	-0.024*	-0.021	-0.023
	(0.012)	(0.011)	(0.011)	(0.010)	(0.010)	(0.012)	(0.012)	(0.013)
Reading Test MS						0.119***	0.070***	
						(0.006)	(0.004)	
Math Test MS							0.128***	
							(0.007)	
(Math-Reading) Test MS								0.054***
								(0.006)
School FE		✓	✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
EDS				✓	✓	✓	✓	✓
Race					✓	✓	✓	✓
Mean	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
R^2	0.06	0.21	0.21	0.23	0.23	0.25	0.28	0.24
N	27,072	27,072	27,072	27,072	27,072	25,175	25,137	25,137

Note: The dependent variable is the same across all specifications: one when taking an advanced English 3 class, zero otherwise. Each specification includes indicator variables for being a female (F), for being non-proficient (Non-Prof) on the English 2 EOC test; and the interaction of those two variables. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Reading Test MS: standardized middle school reading test score. Math test MS: standardized middle school math test score. (Math-Reading) Test MS: difference between the standardized math and English middle school test scores. School FE: school fixed effects. EDS: Economically Disadvantaged Student. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.

Figure 4: Regression Discontinuity, English

Note: Test scores are normalized such that a score of 0 or more means proficiency. Lines are fitted values from regressing a dependent variable that equals one when taking advanced English 3 on indicators for being female and non-proficient on the English 2 EOC test, the interaction of those two variables, and a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, within ±5 of the cutoff. Standard errors are clustered at test score level. The markers represent the proportion of students by gender that take advanced English 3 within each 1 point bin. *Significant at 10%, **5%, ***1%.

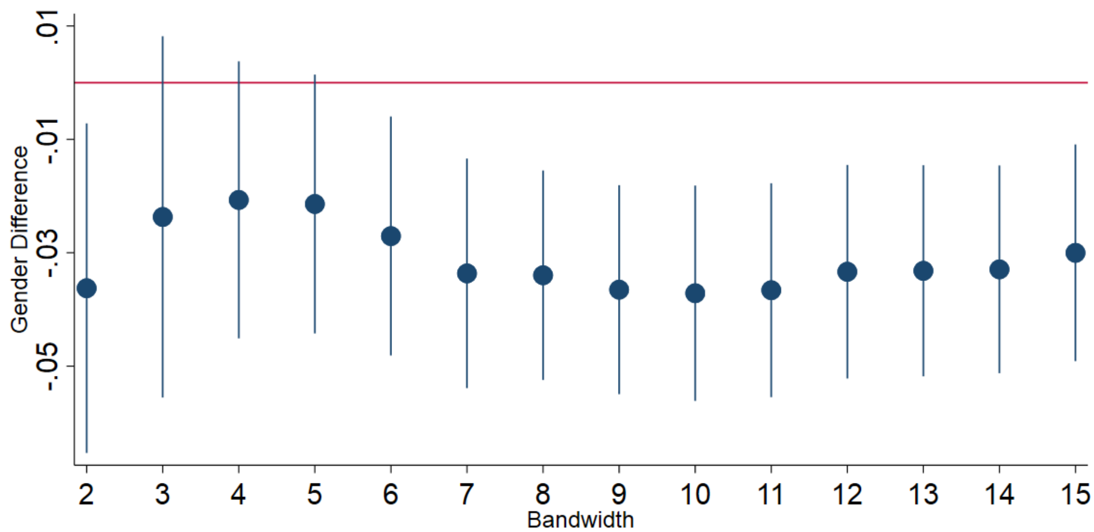
in Figure 4 where there is no discontinuity at the cutoff (zero) for both genders.[19] Once high school fixed effects are included (column (2)) precision improves, which makes the gender difference of about 2 percentage points significant (p-val. < 0.1). This effect does not change in a meaningful way when controlling for year, race and EDS, which is expected given the results from Table 3. The inclusion of middle school test scores, each one individually or the difference between the two, does not change the point estimate much, but the significance changes when controlling for both tests separately.

Figure 5 plots the gender difference in the reaction low performance when estimating the preferred specification that controls for high school and year fixed effects, race, EDS and middle school test scores for different bandwidths.[20] Although not always statistically significant, the coefficient for being a non-proficient women is always negative and relatively stable, between 2 and 4 percentage points, across bandwidths which suggest the existence of a small negative effect that is not precisely estimated for the smaller bandwidths. This implies that women also react differently than men when performing poorly on the English

[19] Figure 4 is the graphical representation of column (1) Table 5

[20] See Figures A7, and A8 for other specifications.

Figure 5: Estimates of the Difference in the Discontinuity for Different Bandwidths, English



Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female in the preferred specification: an indicator for taking advanced English 3 class regressed on indicators for female, being non-proficient on the English 2 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, high school and year FE, indicators for EDS, race and middle school test scores. Standard errors are clustered at test score level. Optimal bandwidth is between 4 and 5.

2 EOC test, however the effect seems smaller than in math. Comparing these estimates to the proportion of women who score within 5 points of the cutoff and take advanced English 3 (47%), this roughly translates the effect of women getting a non-proficient score to a 4 to 8.5% reduction in the likelihood of enrolling in the advanced version of English 3, less than the estimated effect for math.[21]

4.3 Heterogeneity in the Main Effects

The possibility of heterogeneous effects is studied in Appendix Tables A5 and A8 for math and English, respectively. In these tables, the variable of interest, F*Non-Prof., is interacted with various socioeconomic and academic performance variables in order to determine if different reactions to low performance by gender vary across groups.

For the math case, there is no evidence that the different reaction to low performance by gender differs across economic status, race or high/low academic performance during middle

[21]Figure 6 plots together the gender difference in the discontinuity for math and English. It suggests that women react more to non-proficiency in math than in English.

school. In the English case, there is evidence that black non-proficient women are more likely to enroll in advanced English 3 than non-proficient non-black women, however men do not react to low performance regardless of their race. Additionally, economically disadvantaged students seem to react differently to low performance than non-disadvantaged students, but the reaction does not differ across genders.

5 End of High School Outcomes

In this section, I discuss the results from estimating model (1) when the outcomes are end of high school measures like the number of advanced classes taken during high school or plans to attend a 4-year college, instead of the earlier decisions examined before. These results provide some evidence of the longer term effects of non-proficiency early on during high school, and how these effects are different across genders. Results are only presented for the preferred specification at their respective optimal bandwidths.

5.1 Math

Table 6 presents the results for different outcomes realized during the last years of high school: the probability of taking at least one higher-level advanced math class, the total number of higher-level advanced math courses taken during high school, and the plans to attend a 4-year college. The set of higher-level advanced math classes includes any senior level honors class, as well as any AP, IB or community college math course.

Women are more likely to take higher-level advanced classes, as the first two columns of Table 6 show. On the extensive margin, non-proficient and proficient men are equally likely to take a higher-level advanced math class. However, women react differently than men to non-proficiency: the likelihood of ever taking higher-level advanced classes is 4 percentage points lower among non-proficient women.

On the intensive margin, men right below the proficiency cutoff take more higher-level

Table 6: Effect of Non-Proficiency on Math 1 Test on Different Outcomes.

	Prob. higher-level advanced math class	Number of higher-level advanced math classes	4-year College Plan
	(1)	(2)	(3)
Female (F)	0.108***	0.165***	0.141***
	(0.007)	(0.006)	(0.011)
Non-Prof.	0.016	0.045**	-0.002
	(0.010)	(0.016)	(0.015)
F*Non-Prof.	-0.039***	-0.064***	-0.029**
	(0.010)	(0.013)	(0.011)
Mean	0.35	0.47	0.40
R^2	0.27	0.34	0.20
N	13,861	12,004	10,955

Note: Dependent variable indicated at the top of each column. Prob. higher-level advanced math class: one if the student ever took a higher-level advanced math class, zero otherwise. 4-year College Plan: one if student plans to attend a 4-year college, zero otherwise. Each column follows the same specification: dependent variable regressed on a variable equal one when the student is a woman, a variable equal to one when the student is deemed non-proficient on the Math 1 EOC test, the interaction of those two and a first degree polymonial of the test score with flexible slopes above and below the proficiency cutoff. All columns include year and high school fixed effects, controls for EDS, race, and middle school test scores. EDS: Economically Disadvantaged Student. Each column runs a regression withing the optimal bandwidth given the dependent variable: 5, 6 and 4 for columns (1)-(3), respectively. Standard errors are clustered at test score level. *Significant at 10%, **5%, ***1%.

advanced classes, and the opposite is true for women. Although in general women take more higher-level advanced math classes than men, the number is not the same across proficient and non-proficient women, with the latter group taking less higher-level advanced courses. These gender differences in high school preparation could have implications for college major decisions, because women and men might not be equally prepared to undertake certain majors that require a strong mathematical background (Delaney and Devereux, 2019) like STEM majors or Economics.

Finally, women are more likely to plan to attend a 4-year higher education institution. Even though the probability of planning to attend a 4-year college does not change across male students deemed proficient or non-proficient on the Math 1 EOC test, non-proficient women are less likely to state they will attend a 4-year college than proficient ones. This means that the effect of non-proficiency on a first-year test has different effects across gender even at the end of high school on outcomes as important as college attendance.

5.2 English

Table 7 shows the results for some end of high school outcomes of interest: the likelihood of taking higher-level advanced English classes during high school, the total number of higher-level advanced English classes taken during high school, and plans to attend a 4-year college. The set of higher-level advanced English classes includes any junior or senior honors English class, as well as any AP, IB or community college English course. This set of courses includes more than just the courses considered as possible options for English 3 in the main results.

Women are more likely to take higher-level advanced classes regardless of the subject as shown in the first two columns of Table 7. Similar to higher-level advanced math, on the extensive margin non-proficient and proficient men are equally likely to take a higher-level advanced English class. Given that the coefficient for being a non-proficient women is not statistically different from zero, women do not react differently than men to getting a non-proficient English 2 test score. On the intensive margin, the number of higher-level

Table 7: Effect of Non-Proficiency on English 2 Test on Different Outcomes.

	Prob. of higher-level advanced English class	Number higher-level advanced English classes	4-year College Plan
	(1)	(2)	(3)
Female (F)	0.140***	0.314***	0.084***
	(0.008)	(0.011)	(0.005)
Non-Prof.	0.006	0.014	0.018
	(0.013)	(0.022)	(0.010)
F*Non-Prof.	-0.012	-0.051**	-0.023*
	(0.014)	(0.022)	(0.012)
Mean	0.50	0.92	0.38
R^2	0.27	0.33	0.20
N	24,683	24,683	28,701

Note: Dependent variable indicated at the top of each column. Prob. of higher-level advanced English class: one if the student ever took a higher-level advanced English class, zero otherwise. 4-year College Plan: one if student plans to attend a 4-year college, zero otherwise. Each column follows the same specification: dependent variable regressed on a variable equal one when the student is a woman, a variable equal to one when the student is deemed non-proficient on the English 2 EOC test, the interaction of those two and a first degree polymonial of the test score with flexible slopes above and below the proficiency cutoff. All columns include controls for school and year fixed effects, EDS, race, and middle school test scores. EDS: Economically Disadvantaged Student. Each column runs a regression withing the optimal bandwidth given the dependent variable: 5, 5 and 6 for columns (1)-(3), respectively. Standard errors are clustered at test score level. *Significant at 10%, **5%, ***1%.

advanced English classes taken by non-proficient and proficient men do not differ at the cutoff. However, non-proficient women tend to take a lower number of higher-level advanced English classes than the proficient ones.

As in the math case, women are more likely to state plans to attend a 4-year college. There is no difference in the likelihood of planning to attend a 4-year college between proficient and non-proficient male students. However, women react differently than men when they get a non-proficient score on the English test. Non-proficient women are 2.5 percentage points less likely to state plans of attending a 4-year college than proficient ones. This means that the effect of non-proficiency on a test early during high school has different effects across gender even at the end of high school on an outcomes as important as college attendance, regardless of the subject.

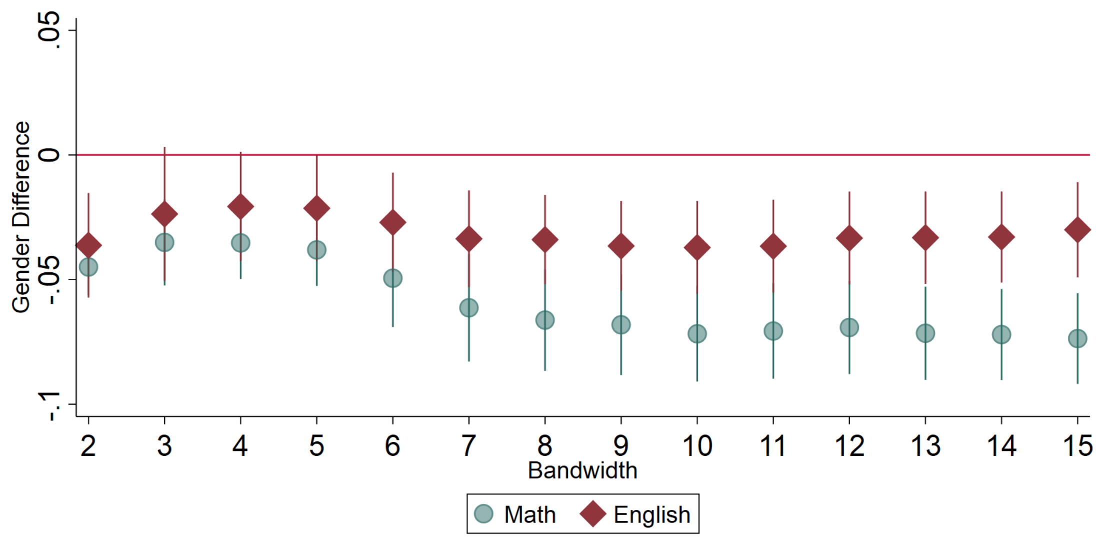
6 Possible Mechanisms

The results from the empirical section only provide evidence about the existence of a gender difference in the reaction to low performance, but they do not give any information about the mechanisms driving the results. However, my findings are consistent with the idea that women interpret their non-proficiency status as a discouraging signal of their ability and this affects their future behavior in terms of class enrollment and college attendance.

Experimental evidence suggests that women are more likely to attribute negative feedback to ability while men tend to attribute it to bad luck (Roberts and Nolem-Hoeksema, 1989; Shastry et al., 2020). Women tend to update their beliefs more pessimistically than men after receiving negative feedback, even when controlling for performance (Berlin and Dargnies, 2016). Moreover, there is experimental evidence that women are not only less willing to compete, but also less likely to do it after losing (Buser and Yuan, 2019), which is consistent with women not taking "harder" classes after performing poorly on the standardized tests.

Additionally, given that the effects of low-performance are stronger in math than in

Figure 6: Estimates of the Difference in the Discontinuity for Different Bandwidths



Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female in the preferred specification: the indicator for taking advanced English 3 or honors Math 2 class regressed on indicators for female, being non-proficient on the English 2 or Math 1 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, high school and year FE, indicators for EDS, race, and middle school test scores. Standard errors are clustered at test score level.

English (see Figure 6), it seems reasonable that the mechanisms are related to the stereotype associated with the domain. For instance, using a lab experiment, Kiefer and Shih (2006) concludes that women are more sensitive to feedback on tests that measure math ability than verbal ability and the opposite pattern is true for men. Similarly, Coffman et al. (2019) finds that male and female beliefs about performance on a task respond more to feedback when the task is gender congruent than when it is not, i.e. men's beliefs respond more to feedback on male-related tasks and women respond more when the feedback is about female-related tasks. In a college context, Kaganovich et al. (2021) finds that female grade sensitivity is category-specific, women are more sensitive to low grades in STEM and Business majors (male-dominated areas) than in Social Science majors. Also, Ellis et al. (2016) suggest that the high dropout rate of women from STEM majors can be related with women's lack of confidence in their math ability, rather than with actual lack of ability.

Another possibility that cannot be ruled out is that parents and/or teachers react differently to a negative outcome depending on the student's gender, which affects students' enrollment decisions. For example, in an experiment with Norwegian families, Tungodden and Willén (2022) document that parents of boys chose a competitive task for their sons more often than parents of girls. However, this explanation is not as convincing because there is evidence of similar gender differences in behavior after instances of "low performance" in college settings, where the influence of parents on the students decisions is lower (Rask and Tiefenthaler, 2008; Kaganovich et al., 2021; Kugler et al., 2021). Regarding teachers, Gentrup and Rjosk (2018) finds that teachers have higher math achievement expectations for boys than for girls, but higher reading achievement expectations for girls. Additionally, girls' math achievement is more negatively affected by low teacher expectations and benefits less from high expectations than boys' achievement. Lavy and Sand (2018) find that primary school teachers' biased expectations in favor of boys encourage them to enroll in advanced math and science courses during high school, but discourage girls.

However, all these explanations can interact with each other, since parents and teachers expectations/stereotypes can affect students self-confidence and reaction to setbacks, impacting their behavior differently across areas. Therefore, there are several possible mechanisms that could explain the gender difference in reaction to low performance that I find, and further research is required to determine how relevant they are and how they interact with each other.

7 Conclusions

This paper uses a RD design to study the decision of North Carolina public high school students to enroll in advanced math or English classes after learning about their performance on EOC tests. The objective is to determine whether there is a gender difference in the reaction to low performance, and if the reaction depends on the subject. Men are usually associated with math and science while women are associated with languages and literature. Therefore, the idea is to determine if the reaction to low performance varies across subjects with different gender stereotypes.

I find a gender difference in the reaction to low performance. In particular, although women are more likely to enroll in advanced classes, the probability of doing so changes depending on their test score performance, while men's likelihood of taking advanced classes does not change with their proficiency status. I find roughly a reduction of 18 to 24% in the probability of enrolling in honors Math 2 when a woman is not proficient on the Math 1 EOC test, but no reduction for low-performance men. The effect seems to be smaller in English than in math, with an estimated reduction of about 4 to 8.5% in the probability of enrolling in advanced English 3 among women that preformed poorly on the English 2 EOC test, and once again no effect of low performance for men. These results suggest that the reaction to low performance might depend on the area in which the hardship is experienced.

If these findings are extrapolated to a career setting, they could be a plausible explanation for women being under-represented in high-level positions. If women and men react to setbacks in different ways, an early defeat might preclude a women from advancing, especially in a male-dominated area. Although the methodology used here does not provide evidence about the mechanisms driving the results, the literature points towards women interpreting their non-proficiency status as a discouraging signal of their ability which affects their future behavior in terms of class enrollment and college attendance. However, further research is required to investigate this or other mechanisms driving the gender differences in the reaction to low performance.

References

AHN, T., P. ARCIDIACONO, A. HOPSON, AND J. R. THOMAS (2019): "Equilibrium Grade Inflation with Implications for Female Interest in STEM Majors," NBER Working Paper 26556, National Bureau of Economic Research.

AUCEJO, E. AND J. JAMES (2021): "The Path to College Education: The Role of Math and Verbal Skills," *Journal of Political Economy*, 129.

AVERETT, C. P. (1994): "Block Scheduling in North Carolina High Schools." https://files.eric.ed.gov/fulltext/ED406734.pdf.

BERLIN, N. AND M.-P. DARGNIES (2016): "Gender differences in reactions to feedback and willingness to compete," *Journal of Economic Behavior & Organization*, 130, 320–336.

BREDA, T. AND C. NAPP (2019): "Girls' comparative advantage in reading can largely explain the gender gap in math-related fields," *Proceedings of the National Academy of Sciences*, 116, 15435–15440.

BUSER, T. AND H. YUAN (2019): "Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad," *American Economic Journal: Applied Economics*, 11, 225–52.

CALONICO, S., M. D. CATTANEO, M. H. FARRELL, AND R. TITIUNIK (2019): "Regression Discontinuity Designs Using Covariates," *The Review of Economics and Statistics*, 101, 442–451.

COFFMAN, K., M. COLLIS, AND L. KULKARNI (2019): "Stereotypes and Belief Updating." Working Paper No. 19-068, Harvard Business School.

DEE, T. S., W. DOBBIE, B. A. JACOB, AND J. ROCKOFF (2019): "The Causes and Consequences of Test Score Manipulation: Evidence from the New York Regents Examinations," *American Economic Journal: Applied Economics*, 11, 382–423.

Delaney, J. M. and P. J. Devereux (2019): “Understanding gender differences in STEM: Evidence from college applications.” *Economics of Education Review*, 72, 219–238.

Eggers, A. C., R. Freier, V. Grembi, and T. Nannicini (2018): “Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions,” *American Journal of Political Science*, 62, 210–229.

Ellis, J., B. K. Fosdick, and C. Rasmussen (2016): “Women 1.5 Times More Likely to Leave STEM Pipeline after Calculus Compared to Men: Lack of Mathematical Confidence a Potential Culprit,” *PLOS ONE*, 11, 1–14.

Frandsen, B. R. (2017): “Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete,” in *Regression Discontinuity Designs*, Emerald Publishing Ltd, vol. 38 of *Advances in Econometrics*, 281–315.

Galindo-Silva, H., N. H. Somé, and G. Tchuente (2020): “Does Obamacare Care? A Fuzzy Difference-in-Discontinuities Approach,” GLO Discussion Paper Series 666, Global Labor Organization (GLO).

Ganley, C. M. and S. T. Lubienski (2016): “Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations,” *Learning and Individual Differences*, 47, 182–193.

Gentrup, S. and C. Rjosk (2018): “Pygmalion and the gender gap: do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling?” *Educational Research and Evaluation*, 24, 295–323.

Goldin, C. (2015): “Gender and the Undergraduate Economics Major: Notes on the Undergraduate Economics Major at a Highly Selective Liberal Arts College,” .

GRANT THORNTON INTERNATIONAL LTD (2021): "Women in Business 2021: A window of opportunity," .

GREMBI, V., T. NANNICINI, AND U. TROIANO (2016): "Do Fiscal Rules Matter?" *American Economic Journal: Applied Economics*, 8, 1–30.

IMBENS, G. AND K. KALYANARAMAN (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 79, 933–959.

JOENSEN, J. S. AND H. S. NIELSEN (2009): "Is There a Causal Effect of High School Math on Labor Market Outcomes?" *The Journal of Human Resources*, 44, 171–198.

KAGANOVICH, M., M. TAYLOR, AND R. XIAO (2021): "Gender Differences in Persistence in a Field of Study," CESifo Working Paper Series 9087, CESifo.

KIEFER, A. AND M. SHIH (2006): "Gender Differences in Persistence and Attributions in Stereotype Relevant Contexts." *Sex Roles*, 54, 859–868.

KUGLER, A. D., C. H. TINSLEY, AND O. UKHANEVA (2021): "Choice of majors: are women really different from men?" *Economics of Education Review*, 81, 102079.

LAVY, V. AND E. SAND (2018): "On the origins of gender gaps in human capital: Short- and long-term consequences of teachers' biases," *Journal of Public Economics*, 167, 263–279.

LEE, D. S. AND D. CARD (2008): "Regression discontinuity inference with specification error," *Journal of Econometrics*, 142, 655–674.

LEVINE, P. B. AND D. J. ZIMMERMAN (1995): "The Benefit of Additional High-School Math and Science Classes for Young Men and Women," *Journal of Business & Economic Statistics*, 13, 137–149.

LONG, M. C., D. CONGER, AND P. IATAROLA (2012): "Effects of High School Course-Taking on Secondary and Postsecondary Success," *American Educational Research Journal*, 49, 285–322.

Macartney, H., R. McMillan, and U. Petronijevic (2018): "Teacher Performance and Accountability Incentives," NBER Working Paper 24747, National Bureau of Economic Research.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

National Center for Education Statistics (2018): "2015 Survey Questionnaires Results: Students' Views of Mathematics, Reading, and Science (NCES 2018-155)," https://www.nationsreportcard.gov/sq_students_views_2015/.

——— (2019): "2015 Survey Questionnaires Results: Classroom Instructions for Mathematics, Reading, and Science (NCES 2018-147)," https://www.nationsreportcard.gov/sq_classroom/#mathematics.

Nord, C., S. Roey, R. Perkins, M. Lyons, N. Lemanski, J. Brown, and J. Schuknecht (2011): "The Nation's Report Card: America's High School Graduates (NCES 2011-462)," .

North Carolina Department of Public Instruction (2015): "2013–14 School Performance Grades (A–F) for North Carolina Public Schools," https://www.dpi.nc.gov/media/5814/open.

——— (2016a): "North Carolina End-of-Course Assessment of NC Math 1," http://www.ncpublicschools.org/docs/accountability/testing/achievelevels/math1achvlvl16.pdf.

——— (2016b): "The North Carolina Testing Program Technical Report 2012–2015 English Language Arts/Reading Assessments (ELA) End-of-Grade 3–8 and End-of-Course English II," https://www.dpi.nc.gov/media/9619/open.

——— (2016c): "The North Carolina Testing Program Technical Report 2012–2015 Mathematics Assessments End-of-Grade 3–8 and End-of-Course Math I," https://files.nc.gov/dpi/documents/accountability/testing/technotes/mathtechreport1215.pdf.

——— (2019a): "End-of-Grade (EOG)," https://www.dpi.nc.gov/districts-schools/testing-and-school-accountability/state-tests/end-grade-eog/.

——— (2019b): "North Carolina End-of-Course Tests," http://www.ncpublicschools.org/accountability/testing/eoc/.

——— (2022): "Course Code Manual Revised March 2022," https://www.dpi.nc.gov/media/14282/open.

Papay, J. P., R. J. Murnane, and J. B. Willett (2016): "The Impact of Test Score Labels on Human-Capital Investment Decisions," *Journal of Human Resources*, 51, 357–388.

Pew Research Center (2021): "The Data on Women Leaders," .

Rask, K. and J. Tiefenthaler (2008): "The role of grade sensitivity in explaining the gender imbalance in undergraduate economics," *Economics of Education Review*, 27, 676–687.

Riegle-Crumb, C. and M. Humphries (2012): "Exploring Bias in Math Teachers' Perceptions of Students' Ability by Gender and Race/Ethnicity," *Gender & Society*, 26, 290–322.

Roberts, M. R. and T. M. Whited (2013): "Chapter 7 - Endogeneity in Empirical Corporate Finance1," Elsevier, vol. 2 of *Handbook of the Economics of Finance*, 493–572.

Roberts, T.-A. and S. Nolem-Hoeksema (1989): "Sex Differences in Reactions to Evaluative Feedback," *Sex Roles*, 21, 725–747.

Rose, H. and J. R. Betts (2004): "The Effect of High School Courses on Earnings," *The Review of Economics and Statistics*, 86, 497–513.

Shastry, G. K., O. Shurchkov, and L. L. Xia (2020): "Luck or skill: How women and men react to noisy feedback," *Journal of Behavioral and Experimental Economics*, 88, 101592.

Shettle, C., S. Roey, J. Mordica, R. Perkins, C. Nord, J. Teodorovic, J. Brown, M. Lyons, C. Averett, and D. Kastberg (2007): "The Nation's Report Card: America's High School Graduates (NCES 2007-467)," .

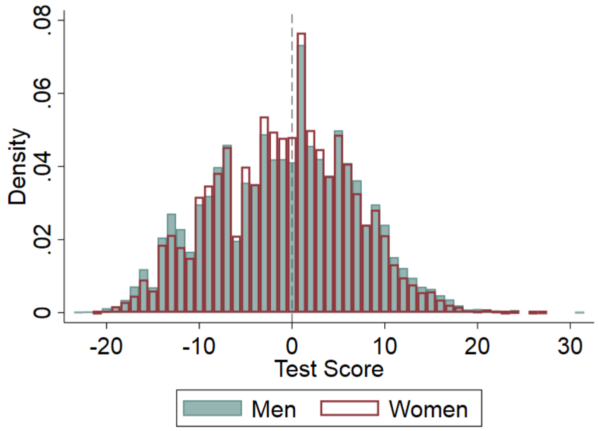
Tan, B. (2020): "Grades as Noisy Signals," *Pedagogy eJournal*.

Tungodden, J. and A. Willén (2022): "When Parents Decide: Gender Differences in Competitiveness," CESifo Working Paper Series 9516, CESifo.

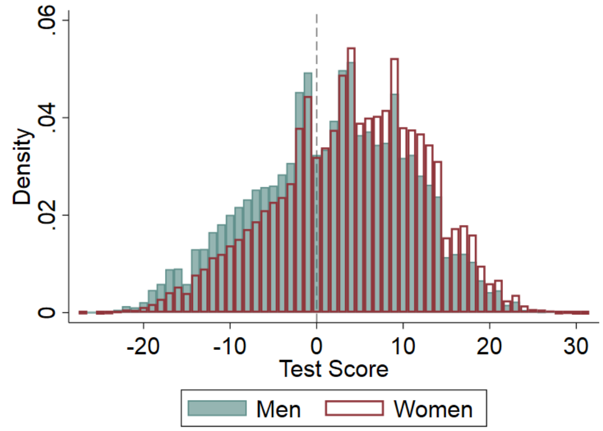
A Appendix

Figure A1: Test Scores Distributions

(a) Math 1.



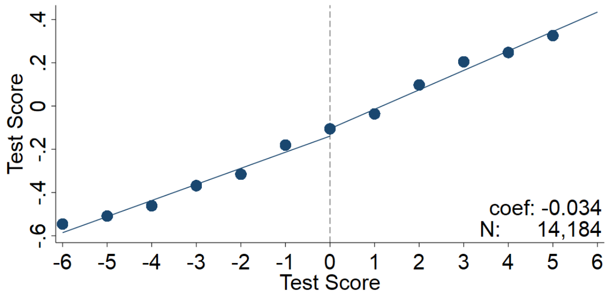
(b) English 2.



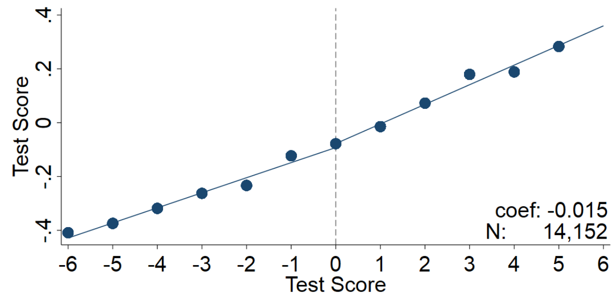
Note: Scores are normalized such that a score of 0 or more means proficiency.

Figure A2: Continuity of Covariates for Math 1 Test

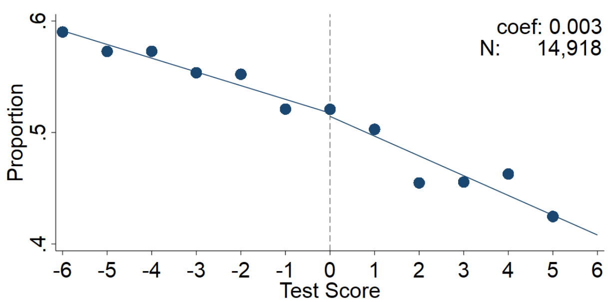
(a) Middle School Math Test Score



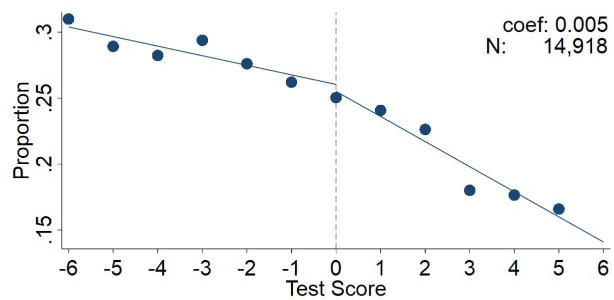
(b) Middle School Reading Test Score



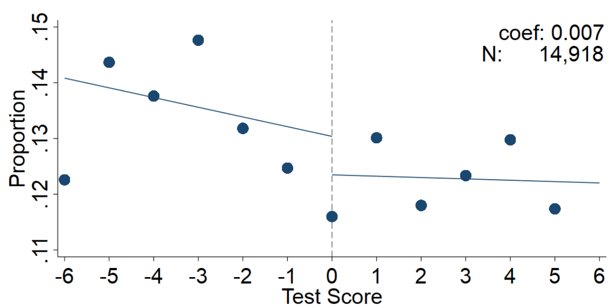
(c) EDS



(d) Black

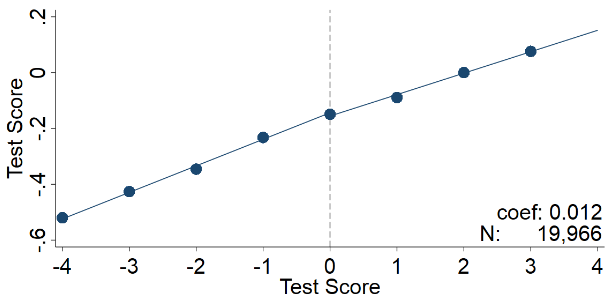


(e) Hispanic

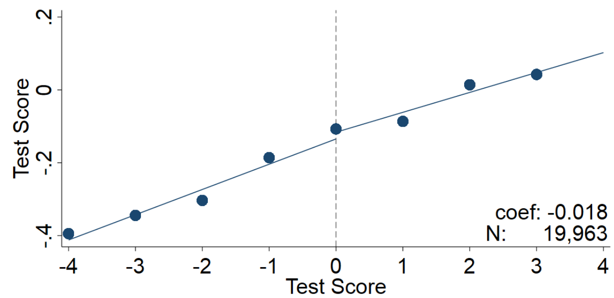


Note: Math test scores are normalized such that a score of 0 or more means proficiency. Lines are fitted values from regressing the corresponding covariate on an indicator for being non-proficient on the Math 1 test. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, within ±6 of the cutoff. Standard errors are clustered at test score level. Middle school test scores are standardized by year. The dots are averages within each 1 point bin. Minority: black or hispanic students. EDS: Economically Disadvantaged Student. Coef: estimated discontinuity, i.e. the coefficient for non-proficiency. *Significant at 10%, **5%, ***1%.

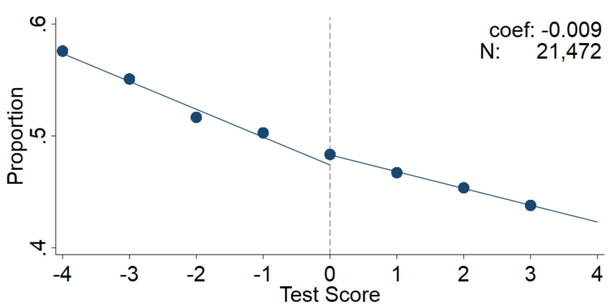
Figure A3: Continuity of Covariates for English 2 Test



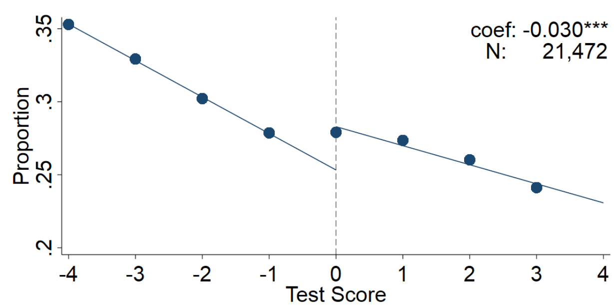
(a) Middle School Reading Test Score



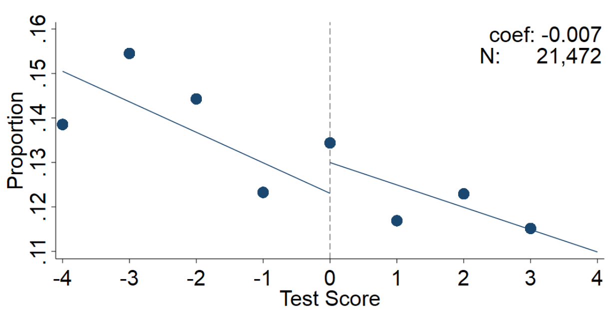
(b) Middle School Math Test Score



(c) EDS



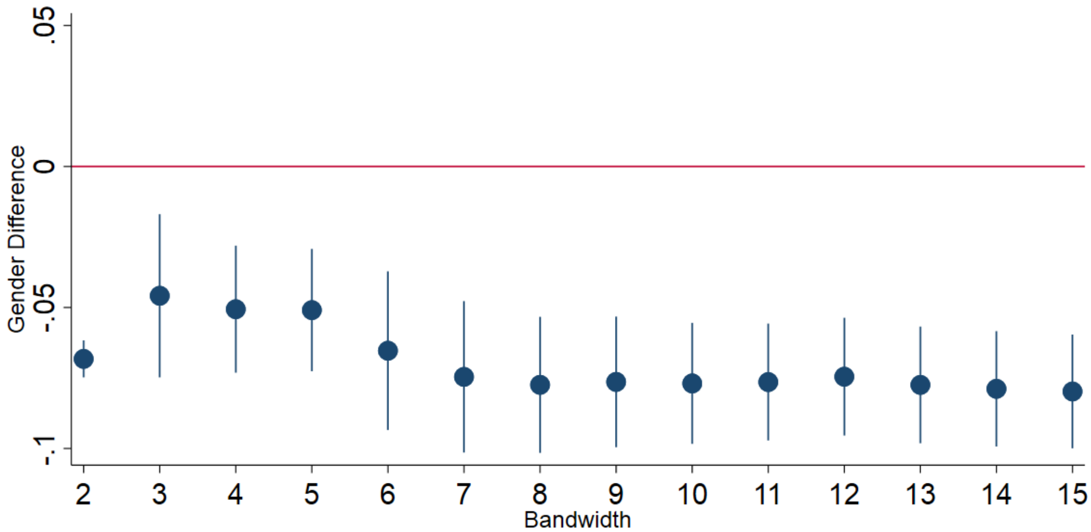
(d) Black



(e) Hispanic

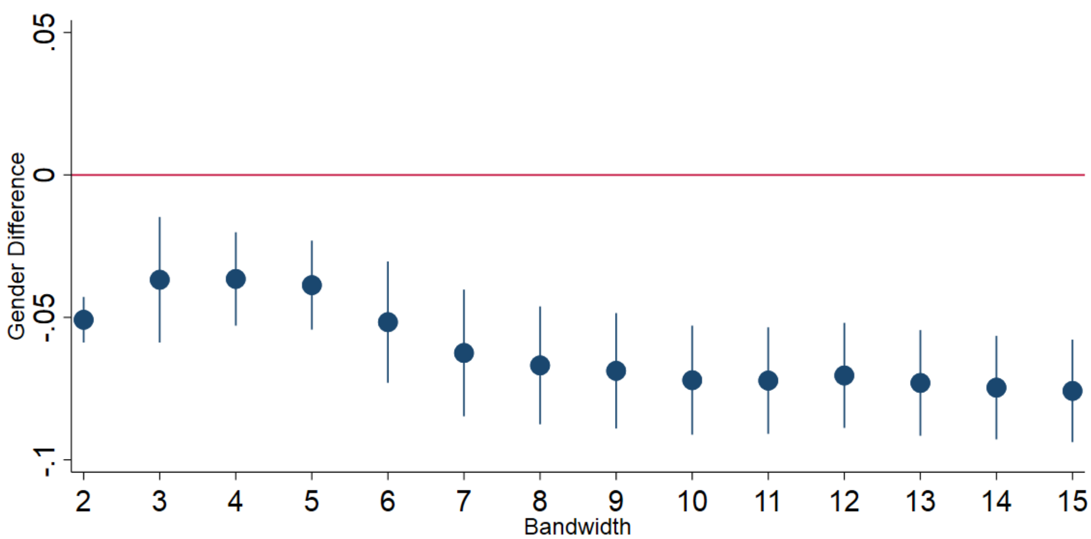
Note: English test scores are normalized such that a score of 0 or more means proficiency. Lines are fitted values from regressing the corresponding covariate on an indicator for being non-proficient on the English 2 test. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, within ±4 of the cutoff. Standard errors are clustered at test score level. Middle school test scores are standardized by year. The dots are averages within each 1 point bin. Minority: black or hispanic students. EDS: Economically Disadvantaged Student. Coef: estimated discontinuity, i.e. the coefficient for non-proficiency. *Significant at 10%, **5%, ***1%.

Figure A4: Estimates of the Difference in the Discontinuity for Different Bandwidths, Math



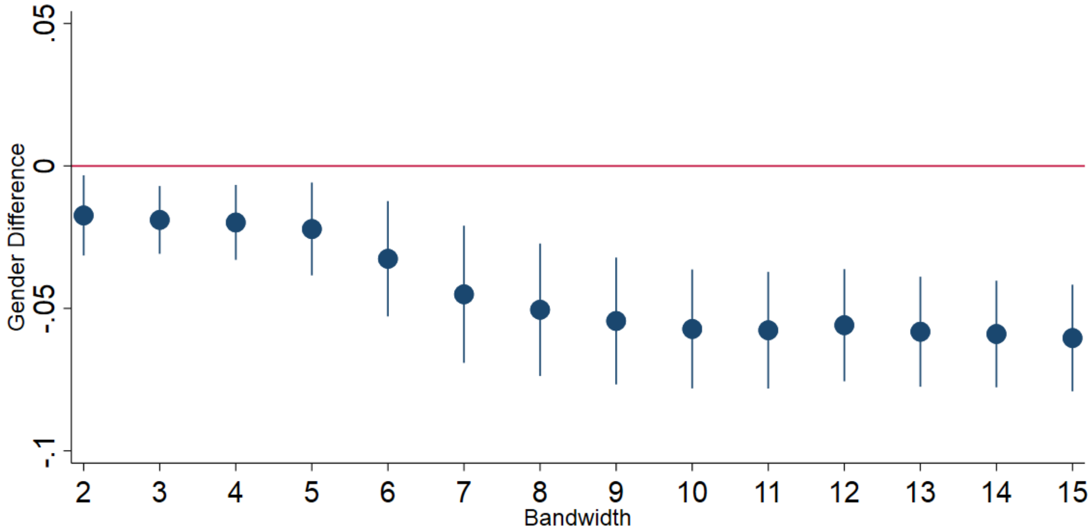
Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female from a regression of an indicator for taking honors Math 2 on indicators for female and being non-proficient on the Math 1 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Standard errors are clustered at test score level. Optimal bandwidth is between 6 and 7.

Figure A5: Estimates of the Difference in the Discontinuity for Different Bandwidths, Math



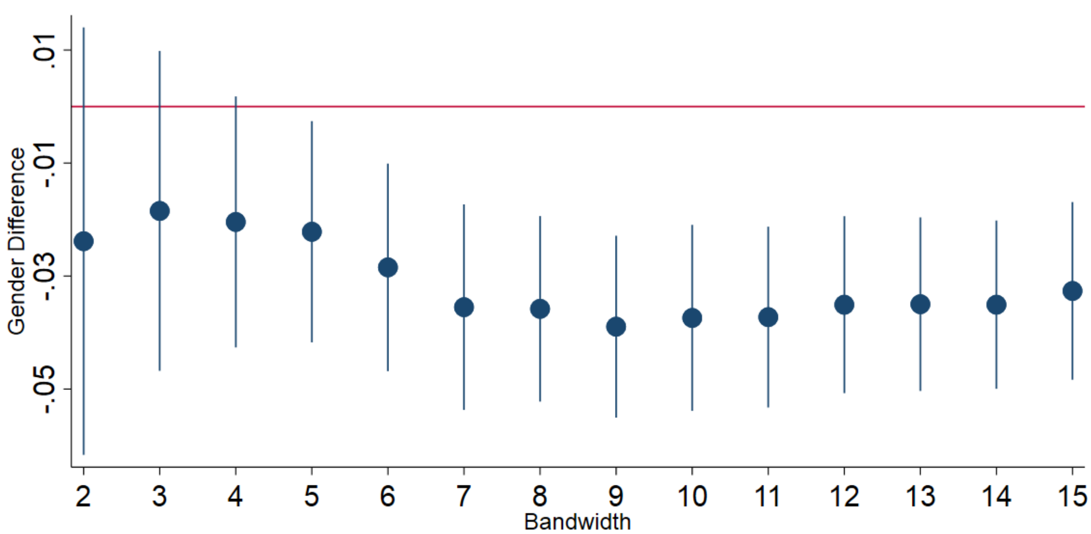
Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female from a regression of an indicator for taking honors Math 2 on indicators for female and being non-proficient on the Math 1 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, year and high school FE, indicators for EDS and race. Standard errors are clustered at test score level. Optimal bandwidth is between 6 and 7.

Figure A6: Estimates of the Difference in the Discontinuity for Different Bandwidths, Math



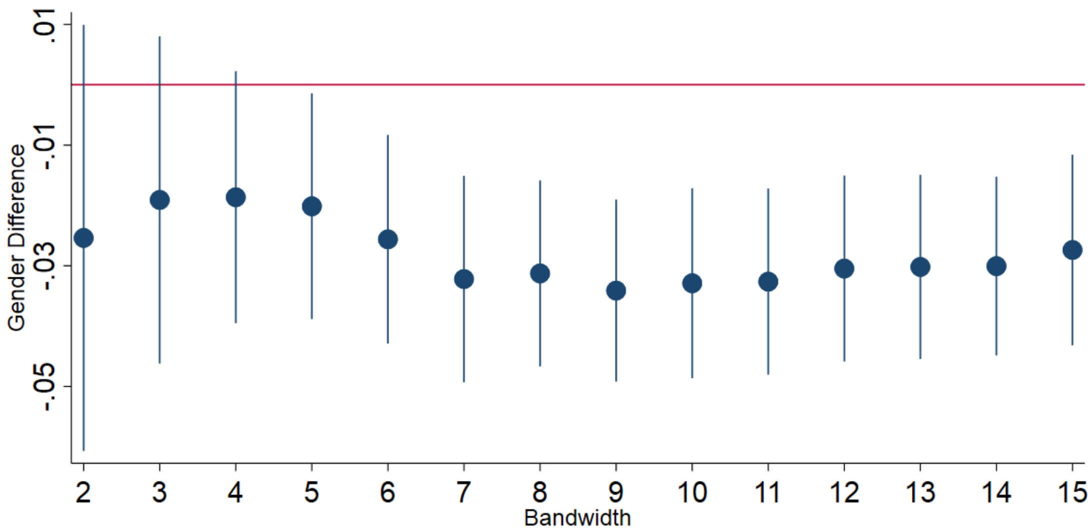
Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female from a regression of an indicator for taking honors Math 2 on indicators for female and being non-proficient on the Math 1 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, year and high school FE, indicators for EDS, race, and the difference between the standardized math and reading 8th grade test scores. Standard errors are clustered at test score level. Optimal bandwidth is between 6 and 7.

Figure A7: Estimates of the Difference in the Discontinuity for Different Bandwidths, English



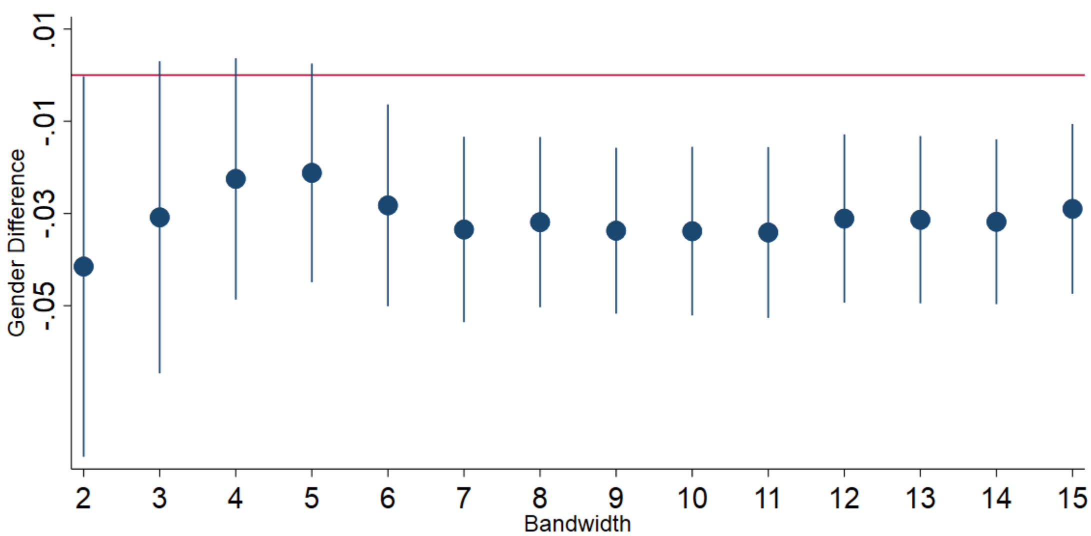
Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female from a regression of an indicator for taking advanced English 3 class on indicators for female, being non-proficient on the English 2 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Standard errors are clustered at test score level. Optimal bandwidth is between 4 and 5.

Figure A8: Estimates of the Difference in the Discontinuity for Different Bandwidths, English



Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female from a regression of an indicator for taking advanced English 3 class on indicators for female, being non-proficient on the English 2 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, high school and year FE, indicators for EDS and race. Standard errors are clustered at test score level. Optimal bandwidth is between 4 and 5.

Figure A9: Estimates of the Difference in the Discontinuity for Different Bandwidths, English



Note: Spikes represent 90% confidence intervals. Dots are the coefficients for being a non-proficient female from a regression of an indicator for taking advanced English 3 class on indicators for female, being non-proficient on the English 2 EOC test, the interaction of those two variables, a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold, high school and year FE, indicators for EDS, race, and the difference between the standardized math and reading 8th grade test scores. Standard errors are clustered at test score level. Optimal bandwidth is between 4 and 5.

Table A1: Comparison Between Fall, After Fall and Middle School Samples

	Fall	After Fall (AF)	Difference Fall - AF	Middle School (MS)	Difference Fall - MS
	(1)	(2)	(3)	(4)	(5)
Female	0.51	0.47	0.03***	0.51	-0.00
Black	0.26	0.33	-0.07***	0.18	0.08***
Hispanic	0.13	0.15	-0.03***	0.10	0.03***
White	0.54	0.45	0.10***	0.64	-0.10***
EDS	0.51	0.59	-0.08***	0.31	0.20***
Non-Prof. Math 1 test	0.49	0.62	-0.13***	0.24	0.25***
Non-Prof. MS Math test	0.63	0.68	-0.05***	0.23	0.39***
Non-Prof. MS Reading test	0.53	0.57	-0.04***	0.18	0.34***
N	27,997	129,541		76,505	

Note: Table presents sample proportions of variables of interest. Fall sample refers to the students that took the Math 1 class during the fall semester of their freshman year. After Fall sample refers to the students that took Math 1 at any other time after the fall semester of their freshman year. Middle School sample refers to the students that took Math 1 during middle school. The difference columns show the mean difference between the Fall sample and the After Fall sample or Middle School Sample, respectively, for each variable and its significance. Non-Prof.: non-proficiency, MS: middle school, EDS: Economically Disadvantaged Student. *Significant at 10%, **5%, ***1%.

Table A2: Comparison Between Fall and After Fall Samples

	Fall	After Fall	Difference
Female	0.51	0.49	0.02***
Black	0.24	0.27	-0.03***
Hispanic	0.11	0.12	-0.01***
White	0.58	0.52	0.05***
EDS	0.42	0.46	-0.04***
Non-Prof. English 2 test	0.35	0.34	0.01**
Non-Prof. Middle School Math test	0.35	0.32	0.03***
Non-Prof. Middle School Reading test	0.38	0.35	0.03***
N	72,395	143,057	

Note: Table presents sample proportions of variables of interest. Fall sample refers to the students that took the English 2 class during the fall semester of their sophomore year. After Fall sample refers to the students that took English 2 at any other time after the fall semester of their sophomore year. The difference column shows the mean difference between the two samples for each variable and its significance. Non-Prof.: non-proficiency, EDS: Economically Disadvantaged Student. *Significant at 10%, **5%, ***1%.

Table A3: Probability of Taking Honors Math 2. ±5 bandwidth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female (F)	0.107***	0.099***	0.099***	0.099***	0.098***	0.099***	0.098***	0.101***
	(0.009)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Non-Prof.	-0.009	-0.006	-0.008	-0.008	-0.008	-0.009	-0.009	-0.009
	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
F*Non-Prof.	-0.051***	-0.041***	-0.040***	-0.039***	-0.039***	-0.038***	-0.038***	-0.038***
	(0.012)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)	(0.008)
Math Test MS						0.099***	0.097***	
						(0.009)	(0.008)	
Reading Test MS							0.006	
							(0.007)	
(Math-Reading) Test MS								0.026***
								(0.005)
School FE		✓	✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
EDS				✓	✓	✓	✓	✓
Race					✓	✓	✓	✓
Mean	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22
R^2	0.05	0.22	0.22	0.23	0.23	0.24	0.24	0.23
N	12,967	12,967	12,967	12,967	12,967	12,322	12,287	12,287

Note: The dependent variable is the same across all specifications: one when taking honors Math 2, zero otherwise. Each specification includes indicator variables for being a female (F), for being non-proficient (Non-Prof) on the Math 1 EOC test, and the interaction of those two variables. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Reading Test MS: standardized middle school reading test score. Math test MS: standardized middle school math test score. (Math-Reading) Test MS: difference between the standardized math and reading middle school test scores. School FE: school fixed effects. EDS: Economically Disadvantaged Student. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.

Table A4: Probability of Taking Honors Math 2. ±7 bandwidth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female (F)	0.125***	0.115***	0.115***	0.115***	0.114***	0.115***	0.114***	0.117***
	(0.013)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)
Non-Prof.	0.001	-0.002	-0.001	-0.001	-0.002	-0.000	-0.000	-0.002
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.008)
F*Non-Prof.	-0.075***	-0.064***	-0.063***	-0.063***	-0.062***	-0.061***	-0.061***	-0.062***
	(0.015)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)
Math Test MS						0.097***	0.095***	
						(0.010)	(0.010)	
Reading Test MS							0.006	
							(0.006)	
(Math-Reading) Test MS								0.025***
								(0.005)
School FE		✓	✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
EDS				✓	✓	✓	✓	✓
Race					✓	✓	✓	✓
Mean	0.23	0.23	0.23	0.23	0.23	0.24	0.24	0.24
R^2	0.08	0.23	0.23	0.24	0.24	0.25	0.25	0.24
N	17,339	17,339	17,339	17,339	17,339	16,465	16,417	16,417

Note: The dependent variable is the same across all specifications: one when taking honors Math 2, zero otherwise. Each specification includes indicator variables for being a female (F), for being non-proficient (Non-Prof) on the Math 1 EOC test, and the interaction of those two variables. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Reading Test MS: standardized middle school reading test score. Math test MS: standardized middle school math test score. (Math-Reading) Test MS: difference between the standardized math and reading middle school test scores. School FE: school fixed effects. EDS: Economically Disadvantaged Student. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.

Table A5: Probability of Taking Honors Math 2.

	(1)	(2)	(3)	(4)	(5)
Female (F)	0.111***	0.113***	0.106***	0.110***	0.105***
	(0.017)	(0.010)	(0.009)	(0.011)	(0.008)
Non-Prof.	-0.014	-0.011	-0.005	-0.011*	-0.008
	(0.008)	(0.006)	(0.006)	(0.006)	(0.005)
Z	-0.064***	0.029	0.040***	0.038	-0.001
	(0.006)	(0.019)	(0.008)	(0.043)	(0.040)
F*Non-Prof.	-0.050**	-0.049***	-0.048***	-0.045***	-0.048***
	(0.021)	(0.013)	(0.010)	(0.013)	(0.009)
F*Z	-0.007	-0.023	0.019	-0.077	0.057
	(0.016)	(0.016)	(0.023)	(0.061)	(0.051)
Non-Prof.*Z	0.013	0.011	-0.016	0.041	-0.074
	(0.009)	(0.016)	(0.014)	(0.051)	(0.095)
F*Non-Prof.*Z	0.002	0.002	-0.010	0.021	0.089
	(0.025)	(0.020)	(0.039)	(0.071)	(0.143)
Z	EDS	Black	Hispanic	Bottom 10% in Middle School Math Test	Top 10% in Middle School Math Test
Mean	0.23	0.23	0.23	0.23	0.23
R^2	0.25	0.25	0.25	0.25	0.25
N	14,184	14,184	14,184	14,184	14,184

Note: The dependent variable is the same across all specifications: one when taking honors math 2 honors, zero otherwise. Each specification includes a dummy variable for being a female (F), for being non-proficient (Non-Prof) in the Math 1 EOC test and a variable Z that changes across columns as indicated by the row Z in the second section of table, all the possible interactions between these three variables and a triple interaction. Additionally, all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. All columns include controls for school and year fixed effects, EDS, race. EDS: Economically Disadvantaged Student. Top (Bottom) 10% refers to a variable equal to one when the student scored in the top (bottom) 10% of the score distribution on their middle school math test. Standard errors clustered at test score level.*Significant at 10%, **5%, ***1%.

Table A6: Probability of Taking Advanced English 3. ±4 bandwidth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female (F)	0.112***	0.100***	0.100***	0.106***	0.107***	0.121***	0.138***	0.115***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Non-Prof.	0.006	0.014	0.014	0.012	0.012	0.011	0.015*	0.013
	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.011)	(0.008)	(0.010)
F*Non-Prof.	-0.015	-0.020	-0.020	-0.018	-0.019	-0.023	-0.021	-0.022
	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.013)	(0.013)	(0.014)
Reading Test MS						0.118***	0.070***	
						(0.006)	(0.005)	
Math Test MS							0.125***	
							(0.005)	
(Math-Reading) Test MS								0.052***
								(0.004)
School FE		✓	✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
EDS				✓	✓	✓	✓	✓
Race					✓	✓	✓	✓
Mean	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
R^2	0.04	0.20	0.20	0.22	0.22	0.24	0.26	0.23
N	21,472	21,472	21,472	21,472	21,472	19,966	19,935	19,935

Note: The dependent variable is the same across all specifications: one when taking an advanced English 3 class, zero otherwise. Each specification includes indicator variables for being a female (F), for being non-proficient (Non-Prof) on the English 2 EOC test; and the interaction of those two variables. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Reading Test MS: standardized middle school reading test score. Math test MS: standardized middle school math test score. (Math-Reading) Test MS: difference between the standardized math and English middle school test scores. School FE: school fixed effects. EDS: Economically Disadvantaged Student. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.

Table A7: Probability of Taking Advanced English 3. ±6 bandwidth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female (F)	0.111***	0.099***	0.099***	0.105***	0.105***	0.118***	0.135***	0.113***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Non-Prof.	0.011	0.012	0.012	0.010	0.011	0.011	0.013**	0.011
	(0.009)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.006)	(0.009)
F*Non-Prof.	-0.026**	-0.028**	-0.028**	-0.025**	-0.026**	-0.029**	-0.027**	-0.029**
	(0.012)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.012)	(0.012)
Reading Test MS						0.116***	0.067***	
						(0.005)	(0.004)	
Math Test MS							0.129***	
							(0.007)	
(Math-Reading) Test MS								0.055***
								(0.006)
School FE		✓	✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
EDS				✓	✓	✓	✓	✓
Race					✓	✓	✓	✓
Mean	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
R^2	0.08	0.22	0.22	0.24	0.24	0.26	0.29	0.25
N	31,495	31,495	31,495	31,495	31,495	29,243	29,199	29,199

Note: The dependent variable is the same across all specifications: one when taking an advanced English 3 class, zero otherwise. Each specification includes indicator variables for being a female (F), for being non-proficient (Non-Prof) on the English 2 EOC test; and the interaction of those two variables. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. Reading Test MS: standardized middle school reading test score. Math test MS: standardized middle school math test score. (Math-Reading) Test MS: difference between the standardized math and English middle school test scores. School FE: school fixed effects. EDS: Economically Disadvantaged Student. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.

Table A8: Probability of Taking Advanced English 3

	(1)	(2)	(3)	(4)	(5)
Female (F)	0.143***	0.125***	0.120***	0.120***	0.118***
	(0.006)	(0.007)	(0.005)	(0.006)	(0.005)
Non-Prof.	-0.015	0.012	0.007	0.007	0.010
	(0.014)	(0.010)	(0.011)	(0.009)	(0.009)
Z	-0.135***	0.017	0.015	0.042	-0.045**
	(0.014)	(0.011)	(0.009)	(0.027)	(0.018)
F*Non-Prof.	-0.026	-0.032**	-0.024*	-0.023	-0.024*
	(0.017)	(0.012)	(0.012)	(0.015)	(0.012)
F*Z	-0.052***	-0.031**	-0.018	-0.060***	0.016
	(0.010)	(0.010)	(0.018)	(0.011)	(0.016)
Non-Prof.*Z	0.048**	-0.015	0.012	0.010	-0.112*
	(0.017)	(0.011)	(0.023)	(0.037)	(0.056)
F*Non-Prof.*Z	0.006	0.032*	-0.003	0.028	0.124
	(0.015)	(0.015)	(0.033)	(0.035)	(0.160)
Z	EDS	Black	Hispanic	Bottom 10% in Middle School Reading Test	Top 10% in Middle School Reading Test
Mean	0.42	0.42	0.42	0.42	0.42
R^2	0.25	0.25	0.25	0.25	0.25
N	25,175	25,175	25,175	25,175	25,175

Note: The dependent variable is the same across all specifications: one when taking an advanced English 3 class, zero otherwise. Each specification includes an indicator variable for being a female (F), for being non-proficient (Non-Prof) on the English 2 EOC test and a variable Z that changes across columns as indicated by the row Z in the second section of table, as well as all the possible interactions between these three variables and a triple interaction. Additionally all specifications include a first-degree polynomial of the running variable with different slopes above and below the proficiency threshold. All columns include controls for school and year fixed effects, EDS, race. EDS: Economically Disadvantaged Student. Top (Bottom) 10% refers to a variable equal to one when the student scored in the top (bottom) 10% of the score distribution in their middle school reading test. Standard errors are clustered at test score level.*Significant at 10%, **5%, ***1%.