Middle School Math Acceleration, College Readiness and Gender: Regression Discontinuity Evidence from Wake County, North Carolina*

Shaun Dougherty
University of Connecticut

Joshua Goodman†
Harvard University

Darryl Hill
Wake County Public Schools

Erica Litke
Harvard University

Lindsay Page
University of Pittsburgh

September 11, 2014

Abstract

To increase the proportion of high school graduates prepared for college-level math, particularly among low income and minority students, the Wake County Public School System implemented a policy assigning middle school students to an accelerated math track on the basis of prior test scores. A regression discontinuity design comparing students just above and below the eligibility threshold shows that acceleration has little effect on standardized test scores. Acceleration has no effect on boys' grades but substantially lowers the grades of girls in recent cohorts. In the earliest cohorts, the vast majority of those accelerated in 8th grade, including minority students, remain on the college-ready track in high school by passing, though not excelling in, geometry in 9th grade. These results suggest that middle school math acceleration has promise for increasing college readiness among disadvantaged populations but that girls' math performance may suffer in settings where they are near the bottom rather than the top of the skill distribution.

*This research was made possible by the broader collaboration between the Wake County Public School System and the Strategic Data Project, housed in Harvard's Center for Education Policy Research. We are grateful for feedback provided by David Deming, Eleanor Dillon, and Maria Fitzpatrick, as well as seminar participants at the University of Connecticut, University of Pittsburgh, Carnegie Mellon, APPAM, AEFP, SREE and AERA. We gratefully acknowledge institutional support from the Taubman Center for State and Local Government at Harvard. Any errors are our own.

†Corresponding author.

1 Introduction

In the US, mathematics achievement is often regarded as essential for individual educational and economic success as well as national global competitiveness (Chazan, 2008; The College Board, 2000). Since Sputnik in 1957 and later A Nation at Risk (1983), policymakers have called for increased proficiency in math among American students as a national imperative (Gardner, 1983; Tate, 1997). Efforts thus have been made to increase the amount and rigor of mathematics course taking, with a focus on exposure to algebra (Adelman, 2006). These efforts have been bolstered by a body of research, albeit descriptive, suggesting that access to algebra is a key determinant of future academic success, given that algebra serves as a gatekeeper to higher-level mathematics (Adelman, 2006; Chazan, 2008; The College Board, 2000; Education Commission of the States, 2008; Ladson-Billings, 1997; Stein, Kaufman, Sherman, & Hillen, 2011). Yet, not all students appear to have equal access to algebra. Traditionally, students are "selected" into Algebra I classes either via their own volition, parental choice or counselor/teacher recommendation (Stein et al, 2011). Furthermore, certain groups of students, particularly black students, Hispanic/Latino students, and students from low-income backgrounds, are under-represented in higher-level courses, thus access to algebra is considered an issue both of equity and of civil rights (Moses & Cobb, 2001).

Driven by national imperatives on mathematics education and concerns about inequitable access to advanced mathematics, hotly debated universal algebra policies have been a common policy response. Such policies aim to increase the shares of students exposed to algebra by mandating that all 9th or even 8th graders enroll in Algebra I (Loveless, 2008; Schneider, 2009; Silver, 1995). These policies hope to increase the numbers of students enrolling in Algebra I in order to allow more students the opportunity to succeed in more advanced courses while in high school. In many cases, such policies are also intended to mitigate equity concerns - promoting access for students traditionally underrepresented in higher level mathematics by removing barriers to entry. While mandating all students to take algebra addresses the advancement and equity concerns on the surface, some have argued that this blunt policy instrument has detrimental effects (Loveless, 2008; Nomi, 2012). In truth, few such programs have been evaluated, and of those that have, the findings are mixed (Allensworth, Nomi, & Montgomery, 2009; Burris, Heubert, & Levin, 2006;

Clotfelter, Ladd, & Vigdor, 2011; Nomi & Allensworth, 2013; Rickles, 2011; Stein et al, 2011).

Concerns around the effectiveness of universal policies focus on student readiness for algebra. Selective but non-systematic entry may deny access to students who are actually prepared to take on algebra coursework, placing unfair barriers to future opportunities in mathematics and exacerbating pre-existing inequalities. However, universal algebra policies may force underprepared students into a rigorous course in which they may not succeed. Thus, "opening the gates" may actually have the opposite effect as intended and depress advanced mathematics course-taking by increasing mathematics failure rates (Gamoran & Hannigan, 2000). Given these concerns, universal algebra policies have fallen out of favor and districts have begun to seek alternative, objective mechanisms to advance students into the mathematics pipeline. It is therefore important to understand the benefits and consequences of alternative mathematics course assignment strategies.

The Wake County Public School System (WCPSS) in Wake County, North Carolina recently responded to similar concerns around advancement in and equitable access to mathematics. WCPSS has attempted to increase advanced math course-taking by ensuring that all students who are prepared to be successful in Algebra I are enrolled in the course as early as possible in their academic trajectory. In contrast to the algebra-for-all policies, WCPSS developed and implemented a targeted enrollment strategy beginning with the 2010-2011 academic year. WCPSS utilizes a proprietary numeric criterion developed by the SAS Institutes Education Value-Added Assessment System (EVAAS) to determine student eligibility for an accelerated math curriculum in grades 6 and 7 and for placement in Algebra I in grade 8. In this paper, we capitalize on this criterion-based assignment policy to examine the impact of accelerated course placement on students' short run academic outcomes. Because the policy was not implemented with strict fidelity – students had the ability to opt out of their recommended course placement – we utilize a regression discontinuity analytic strategy together with instrumental variables to handle the "fuzzy" discontinuity (Imbens & Lemieux, 2008).

As a preview, our analyses yield three main findings. First, math acceleration had no clear effect on standardized test performance in mathematics. Second, we find that acceleration had no effect on the course grades of boys but substantially lowered math course grades for girls. Among students just below the EVAAS threshold, girls earn math grades that are nearly one grade point higher than boys on average. The negative effects of acceleration on girls, however, are more than enough to reduce that difference to zero, so that accelerated girls earn grades similar to boys. On average, whereas girls would otherwise earn As or Bs, those who just met the criterion for acceleration instead earned Cs or below. These negative grade results for girls are driven by the most recent cohorts. For the oldest cohort, which is old enough to now observe in high school, we find that a large share of students who are accelerated in mathematics in middle school continue on the accelerated track in high school. Together, these results suggest that middle school math acceleration has promise for increasing college readiness, though girls' math performance may suffer in settings where they are near the bottom of the relative skill distribution.

The remainder of the paper is structured as follows. In the section that follows, we highlight to key literature relevant to our investigation. In Section 3, we provide a history and detailed description of the middle-grades math acceleration policy in WCPSS. We then outline the data and empirical strategy for our investigation in Section 4. In Section 5, we present results, and Section 6 concludes with a discussion of these results and implications for policy, practice and future research.

2 Literature Review

With its gatekeeper status is acknowledged, Algebra I enrollment rates have increased. In particular in the past two decades, there has been significant growth at the 8th grade level while algebra enrollment in the 10th and 11th grades has declined, reflecting the push for algebra earlier in students mathematics careers (Stein et al, 2011). Yet, this shift has not been experienced uniformly. For example, among 8th graders, black students and Latino students continue to enroll in Algebra I at lower rates than their white counterparts (Gamoran & Hannigan, 2000; Stein et al, 2011). This may be due to a combination of differences in mathematical under-preparedness and course placement practices. In addition, black students and Latino students may be enrolled disproportionately in schools in which algebra is not offered in the middle school years. Taken together, concerns about the equality of access to algebra courses remain and have been foregrounded in recent years, particularly given evidence of the positive outcomes associated with taking algebra.

Research has documented the effects of algebra enrollment on later course-taking as well as on a variety of educational and economic outcomes (NCTM, 1989; Ham & Walker, 1999, Stein et al, 2011). Taking introductory algebra early in ones educational career is a strong positive predictor of taking subsequent, more advanced mathematics courses (Gamoran & Hannigan, 2000; Stein et al, 2011). Given the sequential nature of secondary mathematics courses, this is particularly important in light of research showing that taking courses beyond Algebra II is a significant predictor of college readiness and persistence (Adelman, 2006). Yet, research also shows that certain students, particularly those who are low-achieving or underprepared, may be negatively impacted by early exposure (Gamoran & Hannigan, 2000; Loveless, 2008; Nomi, 2012).

Importantly, much of the research on the relationship between algebra course taking and student outcomes suffers from an important form of selection bias (Stein et al, 2011) and therefore does not support causal claims about the impact of algebra. Students are not randomly assigned to algebra in 8th or 9th grade. Rather, they typically are assigned based on some combination of teacher or counselor recommendation, prior achievement levels, and student or parent preferences. These factors, as well as unobserved differences between students who take algebra and those who do not in a given grade, may serve to predict subsequent achievement. Therefore, comparing outcomes for those who have and have not taken algebra in 9th grade, for example, will not provide causal evidence on the benefits of taking algebra early in ones secondary school career.

A concern that arises from the typical course selection processes relates to those students who may be overlooked. The selection processes, together with factors such as teacher expectations and school course offerings, can lead to circumstances in which some students who are prepared for a course are excluded nonetheless. Indeed, despite sufficient preparation, certain demographic groups are not proportionally represented in algebra courses (Stein et al, 2011). For example, in one district Stone (1998) finds that even among high achieving students, those from high socioeconomic status (SES) backgrounds were three times as likely to be assigned to gatekeeper courses such as Algebra I, compared to their lower-SES peers.

In response to concerns regarding equity in exposure to algebra, some policymakers have advocated for early and universal access to algebra, with some districts mandating Algebra I for all students in the 9th grade and others for all 8th grade students (Bitter & ODay, 2010; Burris, et al, 2006; The College Board, 2000; Loveless, 1998; Silver, 1995). Yet, the Algebra-for-All movement has generated substantial debate (Loveless, 2008; Schneider, 2009; Silver, 1995), at the heart of which lies a tension regarding student readiness for algebra. Selective entry may deny access to those who are prepared to take the course, placing unfair barriers to future opportunities. Nevertheless, universal algebra policies may force underprepared students into a course in which they may not be successful, particularly without additional supports. Thus, opening the gates may exacerbate pre-existing inequalities by increasing mathematics failure rates and lowering grade point averages (Gamoran & Hannigan, 2000).

Students' academic outcomes are, in part, predicated on the program of coursework they follow in high school. If they begin high school in remedial math, there is a ceiling on the number and types of college-preparatory math courses they can take. Universal policies address academic outcomes in theory by interrupting the selectivity of enrollment in college-preparatory courses. Students no longer risk being placed off the college-preparatory track based on background characteristics or prior achievement. If students have access to Algebra I in the 9th grade (and pass the course), then they have the opportunity to progress in mathematics beyond Algebra II and to obtain credits in advanced mathematics during high school (Allensworth, et al, 2009).

While selective assignment undoubtedly excludes students who are prepared to take algebra, universal policies force students who may not be prepared mathematically into algebra classes. This may be detrimental to both groups of students (Loveless, 2008). In an evaluation of algebra assignment policies in Charlotte-Mecklenburg (NC), Clotfelter and colleagues (2011) find negative effects of accelerating low-skilled students into Algebra I in 9th grade. Research on universal algebra in Chicago revealed that the mathematics achievement of high-skilled students declines in heterogeneous classes as a result of the district's policy (Nomi, 2012). Additionally, universal policies may obscure actions and adjustments by schools and teachers. While some schools may adapt pedagogy in rigorous ways to meet the needs of a more heterogeneous population, others may "water down" the curriculum and nevertheless perpetuate previous systems of inequality

(Schneider, 2009). Universal policies also implicitly mandate changes to students' preparation for algebra, having curricular implications for the grades prior to those in which students take algebra. These changes may not occur in practice. Simply mandating that all students take algebra without giving attention to students' preparation, what algebra entails, or how it is taught may be damaging to the very students the policy was intended to help.

Yet, little causal evidence exists on universal algebra policies (Stein et al, 2011). Across existing studies, researchers unsurprisingly find that universal policies increase algebra enrollment (Allensworth, et al, 2009; Burris et al, 2006; Everson & Dunham, 1996; Stein et al, 2011). Impacts on student achievement, however, are mixed (Stein et al, 2011). Clotfelter and colleagues (2011) and Nomi (2012) find negative impacts of algebra acceleration policies. In Chicago, providing an algebra curriculum for all 9th graders increased algebra credit accumulation but also failure rates across ability groups (Allensworth et al, 2009). Further, universal algebra did not raise standardized test scores, though dropout rates remained stable. Nomi and Allensworth (2009) do, however, find positive short-term impacts on GPA and standardized test scores for lower performing students when required to take an additional period of algebra (e.g., a double dose of algebra). A positive and significant increase in course failure rates is nevertheless also a consequence. Cortes, Goodman and Nomi (2012) find that in the longer term, this same double-dose strategy has positive effects on ACT performance, high school graduation, and college entrance. Thus, there may be promise for algebra enrollment policies when combined with appropriate support for under-prepared students, although care must be taken with how such policies are implemented (Nomi & Allensworth, 2013).

In sum, the limited extant research suggests a need for more empirical work to evaluate the causal effects of algebra assignment policies and to identify policies that might best encourage early and equitable exposure for students who are prepared. As policymakers consider existing studies on the impact of universal algebra, they may seek to avoid the pitfalls associated with such policies by attempting to identify students who are likely to be successful in algebra course work. Through the proposed study, we will add to the body of empirical evidence on the impacts of early algebra and accelerated mathematics by examining the effects of the WCPSS initiative on subsequent course-taking, mathematics achievement and other student-level outcomes.

3 Math Acceleration in Wake County

District leaders in WCPSS initiated the targeted enrollment policy to respond to two key concerns. First, approximately 30 percent of WCPSS 8th graders enrolled in Algebra I and district leaders hoped to increase the overall enrollment. Second, the district had concerns that the students who did enroll in Algebra I in the 8th grade were not demographically representative of the district overall. In response, the school board, partnering with a task force focused on the experiences of economically disadvantaged students, sought a strategy to provide equitable access to appropriate and rigorous mathematics courses in the middle grades and to ensure access to Algebra I by the 8th grade for academically prepared students. In particular, the district hoped to increase the disproportionately low rates of enrollment in accelerated math coursework among black students, Hispanic students, and students from low-income households. The district's theory of action assumed that increasing students access to such coursework prior to high school would, in turn, increase their subsequent academic opportunities and, specifically, their likelihood of completing a rigorous, college-preparatory sequence of high school math courses.

The district ultimately implemented a targeted middle-grades math enrollment strategy. Starting in the 2010-11 school year, the district identified students eligible for accelerated math using a proprietary numeric criterion developed by the SAS Institutes Education Value-Added Assessment System (EVAAS). At the end of each academic year, the EVAAS model generates for each student a predicted probability of success on the North Carolina Algebra I end-of-course exam, based on all available prior standardized end-of-grade test scores.[1] The district stipulated that students with a 70% or higher probability of Algebra I success would be recommended for placement in accelerated math courses. For 6th graders, such a course was often called "accelerated math", for 7th graders it was pre-algebra, and for 8th graders it was Algebra I.

In the accelerated math course for sixth graders, the course standards include all of the sixth

[1]For purposes of the policy studied here, success is defined as achieving Level III on the Algebra I end-of-course exam. The EVAAS model also generates predicted probabilities for a given student achieving other levels, none of which is relevant here.

grade content for the non-accelerated course and roughly one half of the content for the non-accelerated seventh grade course. Similarly, in Pre-Algebra, the accelerated math course for seventh graders, the standards include the remaining content for the non-accelerated seventh grade course, and all of the content for the non-accelerated eighth grade course. The subject matter of the sixth and seventh grade advanced math courses overlaps largely with the standards that are tested on the North Carolina End-of-Grade examinations, and content review for the End-of-Grade examinations is included within each course outline. The eighth grade advanced course, which is Algebra I, includes the content typically covered in a high school first-year algebra course, but as of the 2012-13 school year, is an integrated course that is part of a three-year high school sequence comprising the material in the Common Core State Standards for Mathematics. As with the other accelerated courses, content review for the eighth grade End-of-Grade math examination is also included as a part of the course outline .

WCPSS leadership worried that an algebra-for-all policy might enroll students in courses for which some were not academically prepared. Use of the EVAAS predicted probability had two perceived advantages. First, it helped identify students who were thought to be well-prepared for such coursework. Second, because EVAAS is an objective measure, the district believed it could help identify students who might otherwise be overlooked as a result of variation in course grading practices and subjective beliefs about which students are capable of success in accelerated math courses. Indeed, in Dougherty et al. (2014) we document how this new assignment rule strengthened the relationship between academic skill and math acceleration rates while reducing dramatically the role of income and race in course assignment.[2] Interestingly, prior to the new policy, there was little consistent evidence of a gender gap in acceleration rates conditional on academic skill.

The new EVAAS score-based assignment rule thus succeeded in reducing the role of income and race in the math acceleration decision by emphasizing the role of academic skill. The rule also succeeded in increasing overall enrollment in accelerated mathematics coursework in the

[2] In untreated cohorts, low income students spent over 10 percentage points fewer of their middle school years in accelerated math coursework than did their non-low-income peers in the same school and of the same skill, as measured by EVAAS. In the most recent cohorts, this income gap had dropped to three percentage points.

middle grades, and importantly, increasing enrollment for students who were under-represented in such courses. Upon implementation, the policy immediately increased rates of enrollment in middle school accelerated math, and placement recommendations based on this policy have been followed with a high and increasing degree of fidelity.[3] As Figure 1 shows, the share of middle school students in accelerated math rose from 40 percent to nearly 70 percent within two years of the policy's implementation. By 2012-13, nearly all EVAAS-eligible students were enrolled in accelerated math, while acceleration rates remained largely unchanged for students deemed ineligible by the new policy. Acceleration rates rose substantially for both low income and non-low income students though a large income gap in acceleration persists in part because of the large income gap in EVAAS scores. A similar pattern is seen when comparing black and Hispanic students to white and Asian students. Both levels and trends in math acceleration look quite similar for boys and girls.[4] We now turn attention to our analysis of the causal impacts of math acceleration on student outcomes.

4 Data and Empirical Strategy

4.1 Data and Summary Statistics

Using data from the WCPSS longitudinal student information system, we follow students from the end of fifth grade, when they are assigned the EVAAS scores used to determine initial middle school math placement, through middle and high school, during which our outcomes of interest are measured. We can track students as long as they stay within WCPSS. The data include student-level EVAAS scores, generated annually for rising 6th, 7th and 8th graders as further standardized test scores are incorporated into the calculation. The data also contain information on student demographics, such as gender, free/reduced price lunch status, and race/ethnicity. We utilize such variables as controls in some regression specifications and to explore heterogeneity in program impacts.

We observe each student's complete middle school coursework transcript, as well as high

[3] A more detailed discussion of these trends can be found in Dougherty et al. (2014).
[4] See Figures A.2, A.3 and A.1 for trends by income, race and gender.

school transcripts for our earliest cohorts. We can observe the math courses in which students enroll and thus their acceleration status. Because classrooms can be uniquely identified and linked to both students and teachers, we can construct measures of peer composition, such as class size or average prior achievement, and teacher characteristics, such as years of experience or value-added. These classroom-level measures help us characterize in greater detail the various channels through which acceleration may affect student outcomes. We observe three important categories of outcomes that may be affected by math acceleration, namely standardized test scores, grades earned in middle school courses, and the high school coursework in which students later enroll. Standardized test scores come from North Carolinas end-of-grade (EOG) exams in math and reading comprehension, administered in the 3rd through 8th grades regardless of the specific courses in which the students were enrolled. That all students in a given grade receive a common assessment allows us to explore whether acceleration affected math and reading achievement at the end of 6th, 7th and 8th grade.

Because the acceleration policy under study was first implemented in the 2010-11 school year, we focus on data for the 2010-11, 2011-12 and 2012-13 school years. Our main analysis sample consists of WCPSS students with valid EVAAS scores who entered 6th grade in the 2009-10 through 2012-13 school years. We refer to these students collectively as the 2010-13 cohorts, named for the spring of the academic year in which they first entered 6th grade. The 2010 cohort was subject to the new policy starting only in 7th grade, while the subsequent three cohorts were subject to it starting in 6th grade.

Table 1 contains summary statistics for the main analytic sample. Here, and in most of our analyses, each observation is a student-year, so that some students are represented up to three times, once each in 6th, 7th and 8th grades.[5] Column 1, which contains all students in the sample, shows that 57% of WCPSS students in these grades are white or Asian and 38% are black or Hispanic. During this time period, 70% of middle school students are in accelerated math coursework, and the average EVAAS predicted probability is more than 10 percentage points higher than the 70% eligibility threshold set by the assignment rule. In fact, that EVAAS threshold represents

[5]Grade retention in middle school is quite rare in WCPSS, so very few students appear more than three times in the data.

roughly the 25th percentile of math skill in the district, so that the accelerated track would contain about 75% of WCPSS students if the acceleration rule were followed exactly. Over 95% of students pass their middle school math courses though fewer than two-thirds earn an A or a B in those courses.

Columns 2 and 3 divide the sample into students in accelerated math courses and those not. Accelerated students are substantially more likely to be white or Asian and less likely to be black or Hispanic. Accelerated students have much higher math skills, whether measured by EVAAS or by their 5th grade math exam z-score, the latter of which suggests a 1.3 standard deviation difference between the average performance of the two groups. Accelerated students' math classes have much more highly skilled peers, are roughly five students larger, and have fewer black or Hispanic peers than do the math classes of non-accelerated students. Accelerated students are six percentage points more likely to pass their math courses and over 30 percentage points more likely to earn an A or a B. The gap in end-of-grade test scores between these two groups of students is quite similar to the fifth grade gap.

4.2 Regression Discontinuity Design

The substantial differences in academic skill and other factors between accelerated and non-accelerated students would severely bias a simple comparison of these two groups' outcomes. To cleanly identify the impact of math acceleration on student standardized test performance, course grades and course-taking outcomes, we exploit the fact that WCPSS chose an EVAAS predicted probability of 70% as the cutoff for assignment to accelerated math coursework. This fact allows us to use a regression discontinuity (RD) design to compare outcomes of students just above and below that threshold, two groups of students who are nearly identical except that the former group was recommended for acceleration while the latter was not. As such, comparison of these two groups near the threshold should yield estimates unbiased by differences in prior academic achievement or other student characteristics. Because EVAAS scores are recalculated after each grade to incorporate new standardized test scores and because math acceleration may affect such scores and thus subsequent EVAAS values, EVAAS scores calculated at the end of 6th and 7th grades may be partly endogenous to the policy itself. We therefore use as a forcing variable each student's EVAAS score as calculated at the end of 5th grade, prior to the point in time when middle school math acceleration could have affected that score.

For the RD approach to yield valid causal inference, subjects must not be able to manipulate the forcing variable. Given that the EVAAS probability is a predicted value based on a proprietary model with multiple inputs, manipulation would be difficult, if not impossible, for three reasons. First, while WCPSS selected the cutoff criteria of 70%, SAS was responsible for generating the probability values, and the underlying model is not made public. Second, the cutoff scores are a function of prior standardized test performance and students likely have neither sufficient technical knowledge of the policy nor sufficient capability to manipulate their own test performance to impact their placement on the continuum of the forcing variable directly on either side of the cutoff. Third, for the earliest cohorts, students sat for standardized tests prior to the development of the prediction model or assignment policy and could not have anticipated it being implemented.

To confirm this reasoning, we examine the integrity of the forcing variable graphically. Figure 2 shows a histogram of the forcing variable for all students in the main analysis sample, with Panel A showing the full sample and Panel B showing the sub-sample on which our RD analysis will focus. The threshold value of 70% is marked with a vertical dashed line. We observe no discrete change in the density at the threshold, suggesting no obvious manipulation of the EVAAS scores. Though this figure presents the distribution for students across all grades and school years, tests and figures disaggregated by grade and school year look similarly smooth.

The reduced form version of our RD design uses a local linear regression to fit the following model for student i in cohort c, grade g and initial middle school s:

$$Y_{icgs} = \beta_0 + \beta_1 Elig_{icgs} + \beta_2 EVAAS_{icgs} + \beta_3 (Elig * EVAAS)_{icgs} + \mu_{cgs} + \epsilon_{icgs} \tag{1}$$

where *Y* is an outcome such as course grade or test score, *Elig* is an indicator for a student's end of 5th grade EVAAS score exceeding 70%, and *EVAAS* is a students EVAAS score re-centered around that threshold. We include cohort-by-grade-by-school fixed effects so that students are being compared to their peers within the same cohort, grade and school. This improves the precision of our estimates but has little impact on their magnitude, as would be expected given that the threshold is the same throughout the district. The coefficient on the eligibility indicator, β_1, therefore represents the difference in outcomes between students just above and just below the eligibility threshold.

Because compliance with the assignment rule is imperfect, the reduced form does not measure the impact of math acceleration itself but only eligibility for such acceleration. We therefore employ a fuzzy regression discontinuity design (Imbens and Lemieux, 2008) by implementing a two-stage approach to estimate the effect of accelerated math on various outcomes. In the first stage, we use each student's position relative to the EVAAS cutoff as an instrument for time spent in accelerated math courses. The first stage thus takes the form:

$$Accel_{icgs} = \beta_0 + \beta_1 Elig_{icgs} + \beta_2 EVAAS_{icgs} + \beta_3 (Elig * EVAAS)_{icgs} + \mu_{cgs} + \epsilon_{icgs} \tag{2}$$

where the right-hand side variables are defined as above. For estimates of impacts on middle school grades and test scores, we define *Accel* as the fraction of middle school years spent in accelerated math courses up to the point in time when the relevant outcome is measured. The coefficient on the eligibility indicator, β_1, thus represents the difference in the fraction of time spent accelerated between students just above and just below the eligibility threshold. We do this to account for the possibility that acceleration in earlier years has an effect on current outcomes. This definition assumes linearity in the effect of acceleration across years. In a few instances where we are not concerned about prior years' impacts, we define *Accel* as an indicator for contemporaneous enrollment in accelerated math. This choice turns out to make little practical difference to our central estimates.

Our second-stage model takes the form:

$$Y_{icgs} = \beta_0 + \beta_1 Accel_{icgs} + \beta_2 EVAAS_{icgs} + \beta_3 (Elig * EVAAS)_{icgs} + \mu_{cgs} + \epsilon_{icgs} \tag{3}$$

where we replace the potentially endogenous fraction of years accelerated with predicted values from the first stage regression. Our ultimate coefficient of interest, β_1, therefore measures differences in outcomes between those who have spent all of their middle schools year in accelerated coursework and those who have spent none of their years in accelerated coursework. The impact of math acceleration is generated by students whose acceleration status was affected by the eligibility threshold, the compliers in this context (Angrist et al., 1996). Because the exogenous variation is generated by students near the threshold, these estimates represent local average treatment effects for students near the 25th percentile of the math skill distribution. For our primary specification, we will estimate these local linear regressions using a triangular kernel, a bandwidth of 15 EVAAS percentage points, and standard errors clustered by initial middle school. We choose that bandwidth because it is quite close the first-stage and reduced form optimal bandwidths suggested by Imbens and Kalyanaraman (2012). We later show that our results are robust to alternate bandwidths, including the Imbens-Kalyanaraman bandwidth, as well as to the inclusion of demographic covariates as controls.[6]

That inclusion of covariates does not affect our central estimates is unsurprising given that the inability to manipulate the EVAAS score suggests students' demographic characteristics should be balanced across the threshold. We confirm this in Table 2, which tests for discontinuities in demographic characteristics at the threshold by running our first-stage specification with various covariates as outcomes. For the complete sample, shown in panel A, all available covariates appear balanced across the threshold, suggesting that our treatment and control groups look quite similar in terms of race, income, special education and limited English proficiency status, age and gender. The first column, which collapses the data to observation counts within each integer-wide EVAAS bin, shows no discontinuity in the density of observations near the threshold, confirming the visual evidence from Figure 2. Panels B and C divide the sample by gender, showing balanced covariates in the male subsample and largely balanced covariates in the female subsample. The two slight imbalances, girls just above the threshold are slightly less likely to be low income and slightly more likely to have limited English proficiency, point in opposite directions in terms of student disadvantage and thus are likely spurious. We show later that controlling for such covariates has little impact on our central estimates.

[6]Though not shown here, the results presented below are also robust to optimal bandwidths selected by the methods proposed in Ludwig and Miller (2007) and Calonico et al. (ming). They are also robust to the use of a rectangular kernel.

4.3 First Stage Results

To test the usefulness of the EVAAS threshold as a source of exogenous variation in course assignment, we first examine the graphical relationship between students' EVAAS scores and the probability of enrolling in accelerated math coursework. Figure 3 shows this relationship by grade and school year, so that each downward-sloping diagonal represents a single cohort of students. The top row represents the 2010-11 school year during which the new assignment policy was first implemented. The largely untreated 2009 cohort is in the upper right box, as they were 8th graders in the 2010-11 school year, and show little evidence of a discontinuity in acceleration rates near the threshold. The 2010 cohort, who were 7th graders in 2010-11 and 8th graders in 2011-12, show clearer discontinuities, with students just above the threshold noticeably more likely to be accelerated in both 7th and 8th grades compared to those just below the threshold. The subsequent 2011-2013 cohorts, all of whom started middle school under the new assignment rule, show substantial discontinuities in acceleration rates as well. These discontinuities seem particularly striking in 7th grade, though the most recent 2013 cohort shows a substantial discontinuity in 6th grade, the one year for which we can currently observe them. These graphs suggest that the new assignment rule has been more and more faithfully implemented over time and will serve as a strong source of exogenous variation in the probability of a given student being accelerated in math.

Figure 4 shows the actual first stage used below. We define the treatment as the fraction of each student's observed middle school years spent in accelerated math coursework, when we pool the 2010-13 cohorts. The discontinuity here is striking. Students just below the eligibility threshold spend about 35% of their middle school years in accelerated coursework, whereas those just above the threshold spend about 50% of their time in such coursework, on average. We confirm this graphical intuition in Table 3, the first row of which estimates the first-stage regressions described in Equation 2. The remaining rows use as an outcome an indicator for being accelerated in a given grade and year. The first column represents the untreated 2009 cohort, while the second through fifth columns represent the treated 2010-13 cohorts. The final column pools all four treated cohorts.

As expected, there is no evidence that the eligibility threshold affected math acceleration rates in the 2009 cohort's first two years, prior to the policy's introduction. In 8th grade, there is small and only marginally significant evidence of an impact for that cohort in the policy's first year. For the 2010 cohort, which was in 7th grade when the new policy began, eligibility increases the fraction of middle school years spent in accelerated math by 7.3 percentage points. This fraction rises monotonically across subsequent cohorts, so that eligibility increases the fraction of years accelerated by 28.3 percentage points for the 2013 cohort. Pooling all four treated cohorts leads to an average estimated first-stage effect of 13.4 percentage points. For the pooled sample, the F statistic for the eligibility coefficient is nearly 50, well beyond the threshold of 10 suggested by Bound, Jaeger and Baker (1995) for a strong instrument. The second through fourth rows show that much of the strength of this instrument comes from particularly strong impacts in 7th grade acceleration decisions, though smaller effects in 6th and 8th grade are still highly statistically significant.

Changing the bandwidth or controlling for covariates has essentially no effect on this first-stage estimate, as seen in column 1 of Table A.1. Figures A.4, A.5 and A.6 show little clear graphical evidence of first-stage heterogeneity by gender, income or race. Table A.2 estimates such heterogeneity by re-running the first-stage specification with the eligibility indicator interacted with indicators for gender, income and race, as well as the direct effect of the given demographic characteristic. The direct effects, as measured by the non-interacted coefficients, suggest that, conditional on EVAAS scores, acceleration rates do not differ by gender, income or race. There is marginally significant evidence that the threshold has larger acceleration effects on male and minority students, although the differential impact on minorities is not consistent across years. There is no evidence of a differential first-stage by income.

Before turning to the impact of math acceleration on student outcomes, we document a variety of channels possibly responsible for such impacts. The most obvious channel through which math acceleration might affect students is through exposure to a more rigorous curriculum, something we cannot measure beyond our ability to categorize courses based on their titles. We can, however, observe other aspects of the classroom experience to which students are exposed, including the characteristics of the peers and teacher in each student's math classroom. In each classroom, we can characterize the mean and standard deviation of peers' math skills as measured by 5th grade math scores, the total class size, and the fraction of students who are female, low income and black or Hispanic. For many of the students' primary math teachers, we can also identify a value-added measure of that teacher's quality based on prior years' test scores, their years of experience, and their gender.

Figure 5 captures graphically the impact of math acceleration on such channels, showing that acceleration changes substantially the peers to which students are exposed but has little clear impact on the teachers students have. This is confirmed in Table 4, which shows instrumental variables estimates of the effect of contemporaneous math acceleration on contemporaneous classroom characteristics. Relative to non-accelerated students, accelerated students take math with peers who are 1.1 standard deviations higher in math skill. Acceleration does not, however, change the heterogeneity of skill to which students are exposed. Accelerated students end up in classes that are 4.2 students larger and substantially less populated by low income, black or Hispanic students. Acceleration has no impact on the gender composition of one's classmates. There is little evidence that acceleration affects the average quality of students' math teachers though we see suggestive evidence that accelerated students are less likely to be assigned particularly low quality teachers, those with value-added measures one standard deviation below the mean. Acceleration has little impact on the experience level or gender of one's math teacher. These results are not biased by our inability to link some students to teachers, as the final column shows that the probability of such linkages is unaffected by the eligibility threshold.

In total, these results suggest that acceleration exposes students to higher skilled peers and possibly fewer low quality teachers, which might have positive effects, but also to larger class sizes, which might have negative effects. Based on these results, it is worth noting that because of the structure of the policy, students on different sides of the EVAAS threshold had mathematics classroom experiences that differed not only in terms of curriculum and course content but also in terms of the student composition of the classroom itself. In this respect, the treatment that we are assessing with our regression discontinuity design is multi-dimensional in nature and not necessarily the effect of a more advanced mathematics curriculum exclusively.

5 Math Acceleration and Student Outcomes

5.1 Middle School Math Grades and Test Scores

Having established that the eligibility threshold provides a strong source of exogenous variation in the probability of being in the accelerated math track, we now estimate the impact of such acceleration on middle school math grades and test scores. We present two types of evidence, visual evidence of the reduced form relationship between these outcomes and EVAAS scores and instrumental variables estimates of the impact of acceleration on these outcomes.

Figure 6 shows the reduced form relationship between grades earned in middle school math classes and initial EVAAS scores, for EVAAS scores within 15 percentage points of the eligibility threshold. Overall, EVAAS scores have a clear and positive relationship with math grades. There is, however, some visual evidence that students' grades just above the eligibility threshold are lower than would otherwise be predicted by a regression line fitted to points below the threshold. This interpretation is confirmed by Table 5, which presents instrumental variables estimates of this treatment effect on several measures of course performance across different grades and cohorts. The first column of panel A shows the impact of acceleration on the continuous measure of math GPA for all four cohorts in the analysis sample. The coefficient implies that being accelerated lowers a student's GPA by a large and statistically significant 0.5 grade points.[7] This result is driven entirely by the two most recent cohorts, for whom acceleration lowers math GPA by more than one grade point, an effect apparent in both the sixth and seventh grades. Acceleration lowers course passing rates by a marginally significant 11 percentage points and lowers the probability of earning an A or B by a highly statistically significant 29 percentage points. There is no clear or consistent impact of math acceleration on grades earned in non-math courses.

The negative impact on course grades could represent true learning losses or could be the result of math teachers assigning grades based on a curve, given that these treatment effects are estimated off of students induced from classes where they would be near the top of the skill distribution to classes where they are near the bottom. We therefore explore the impact of ac-

[7] The point estimate for this GPA impact is fairly robust to alternate specifications, as seen in column 2 of Table A.1.

celeration on the standardized end-of-grade test scores administered in North Carolina, which do not suffer from this potential confounding factor. Figure 7 shows the reduced form relationship between end-of-grade math test scores and initial EVAAS scores. Unsurprisingly, prior achievement as measured by EVAAS scores has a very clear and positive relationship with subsequent achievement as measured by later test scores. Unlike with course grades, there is no apparent discontinuity in test scores at the eligibility threshold. Point estimates in columns 1-4 of Table 6 confirm this, showing no statistically significant impacts of acceleration on math achievement across any grades or cohorts.[8] There is no evidence, as seen in column 5, that acceleration affects the probability of taking end-of-grade math exams, suggesting that selection bias is not driving these non-results. We also observe no evidence of spillover effects onto other subjects. Reading scores are similarly unaffected by acceleration.

Overall, students induced into the accelerated middle school math track by the new district policy appear not to benefit from such acceleration. As measured by end-of-grade test scores, their achievement is unchanged. Their course grades appear, however, to suffer substantially. Given that accelerated students are, if anything, exposed to more highly skilled peers and teachers, we consider two potential explanations for such negative effects. First, as mentioned above, teachers may simply be grading on a curve, so that acceleration lowers grades by dramatically lowering students' relative ranking in their classrooms' skill distribution. Second, acceleration may discourage students by placing them in larger classes where they are relatively low-skilled when compared to their peers. If some students react badly to such placement, academic performance may suffer even if teachers are not grading on a curve. We turn now to heterogeneity analysis that may help distinguish these hypotheses.

5.2 Heterogeneity by Gender

To explore heterogeneous treatment effects of middle school math acceleration, we show the reduced form relationship between GPA and EVAAS scores among boys and girls in Figure 8. The difference between these panels is striking. Girls just above the eligibility threshold have sub-

[8] The point estimate for this GPA impact is fairly robust to alternate specifications, as seen in column 5 of Table A.1.

stantially lower grades than girls just below it, while no such difference is evident for boys. We estimate these effects in the first column of Table 7, which presents our previous instrumental variable estimates using specifications that interact the instrument and treatment variable with gender indicators. The estimates confirm the visual evidence, suggesting that acceleration lowers girls' GPAs by a highly significant 1.3 grade points and has no significant impact on boys' grades. The difference between the GPA effect on boys and girls is highly statistically significant.[9]

The non-interacted female coefficient implies that, conditional on EVAAS score, girls are earning math grades nearly one grade point higher than boys. The negative effects of acceleration are more than enough to reduce that difference to zero, so that accelerated girls earn grades similar to boys. Some of this drop in girls' grades is driven by a marginally significant 18 percentage point drop in course passing rates, as seen in column 2. Much more is driven by the 62 percentage point drop in the probability of earning at least a B in one's math course, implying that acceleration is turning large numbers of girls who would otherwise earn A's or B's into relatively poor performers. In short, the entire negative effect of math acceleration on GPA is driven by girls. It is also important to note that, like the overall GPA effects documented earlier, these negative effects for girls are driven entirely by the most recent two cohorts and do not appear in the earliest two cohorts for whom we can measure high school outcomes.

Figure 9, which shows standardized test scores by gender, suggests potentially negative effects of acceleration on girls' achievement. The regression estimate in column 4 of Table 7 of the effect on girls is negative 0.16 standard deviations but is statistically insignificant. The impact on boys' achievement is a marginally significant positive 0.36 standard deviations but that result is fairly sensitive to the bandwidth, as seen in column 7 of Table A.1. The p-value at the bottom of the column implies marginally significant evidence that the effect on girls' test scores differs from the effect on boys' test scores. Finally, as with grades, the non-interacted female coefficient suggests that, conditional on EVAAS scores, girls outscore boys in math by 0.26 standard deviations.

No observable demographic characteristic other than gender shows heterogeneous treatment effects. The relationship between EVAAS scores, math grades and math scores by low-income

[9]Table A.1, columns 3 and 4, show versions of these estimates in which the sample has been split by gender. The results are quite similar to the interacted version shown here and are robust to the alternate specifications shown.

status and race are shown in Figures A.7, A.8, A.9, and A.10. We observe no visually obvious discontinuities. Table A.3 confirms that there are no statistically significant differences in grade or score impacts by income or race. The only noteworthy difference in that table is that, conditional on EVAAS scores, low income students have lower grades than their non-low income peers, though this difference does not appear in test scores. There are no such significant differences by race.

Table 8 shows that the heterogeneous impacts observed by gender cannot be explained by any differential nature of the treatment that we can observe. The changes in peer and teacher characteristics induced by acceleration do not vary by gender.[10] We do observe marginally significant evidence that acceleration increases the probability that girls have female math teachers. That effect is, however, not significantly different from the effect on boys. That effect also points in the wrong direction for explaining the negative impact on girls' GPAs, given that existing evidence suggests that having same-sex teachers should, if anything, improve girls' performance (Dee, 2005, 2007). The differential impact of math acceleration by gender thus does not appear to be driven by differences in the treatment itself, unless the treatment varies by gender along important dimensions we cannot observe.

5.3 High School Course Outcomes

We have documented two important facts about the impact of math acceleration on middle school outcomes. First, acceleration has little clear impact on test scores. Second, acceleration has substantial negative effects on girls' math grades, a result driven entirely by the most recent two cohorts in the data. These are important findings but the question of ultimate importance is whether such acceleration succeeds in putting students on a high school math trajectory that improves their college readiness.

We cannot observe high school outcomes for the 2012 and 2013 cohorts driving the gender results highlighted above. We can, however, observe high school freshman course enrollment and grades for the 2010 and 2011 cohorts. A student continuing on the accelerated (i.e. college-

[10]These aspects of the treatment also do not vary by income or race, as seen in Table **??**.

ready) track from middle school should be taking a course equivalent to geometry (or higher) during freshman year of high school. For each student in those earliest two cohorts, we construct a measure indicating enrollment in geometry or a higher level math course. We also construct indicators for whether that student passed such a course and whether that student earned a B or higher in such a course.

To test what happens to the marginal accelerated students once they reach high school, Table 9 limits the sample to those from the 2010 and 2011 cohort who were present in WCPSS in 8th grade. We then regress high school course outcomes on 8th grade acceleration status, where acceleration is instrumented using the eligibility threshold, as before. We are therefore estimating the impact of being accelerated in 8th grade on 9th grade math coursework. These estimates measure the LATE of 8th grade acceleration on 9th grade outcomes, for the marginal student accelerated in 8th grade because of the policy.

The top coefficient in column 1, which contains all students, implies that 8th grade acceleration increases by 81 percentage points the probability of being on the accelerated track in 9th grade and thus enrolling in geometry. The second row implies that geometry passing rates increase by 72 percentage points, meaning that nearly all who take geometry because of the assignment rule end up passing the class. In that sense, acceleration appears to succeed in placing a large fraction of marginal students on the college-ready track in high school. On the other hand, row three shows little evidence that acceleration increases the proportion of 9th graders earning A or B grades in their geometry classes. In short, a high fraction of students induced into accelerated math in middle school continue on that track in high school and nearly all such students end up passing their first such high school math class. Nearly all of them do so, however, by earning Cs and Ds, which may indicate that they will have challenges succeeding in higher level math courses or positively signaling their preparation for college-level mathematics.

The remaining columns of Table 9 separate these results by gender, income and race. Doing so increases the size of the standard errors, so that none of the differences between groups are statistically significant. It also substantially weakens the instrument for many of the listed subgroups, suggesting that the resulting point estimates may be biased toward their OLS counterparts. As such, though the point estimates suggest that a higher fraction of accelerated girls remain accelerated in high school and pass freshman geometry, these magnitudes should be interpreted with caution. We also note that, even if these point estimates are correct, they do not contradict our earlier results concerning the drop in girls' middle school GPA because the earlier results were driven by the 2012 and 2013 cohorts who are not yet old enough to be observed in high school. This raises the possibility that these later cohorts of students, particularly girls, may not benefit in the long-run as much as the earlier cohorts appear to do.

Low income and non-low income students appear to remain on the accelerated track at similar rates, though the separation by income dramatically weakens the eligibility instrument. The instrument remains quite strong, however, for black and Hispanic students, with estimates suggesting that acceleration increases geometry enrollment rates by 85 percentage points and geometry passing rates by 77 percentage points. This is a remarkable improvement over the seven percent passing rate for students just below the eligibility threshold. Nonetheless, there is no evidence of improvement in the rate at which black and Hispanic students earn As or Bs in freshman geometry. The instrument is too weak for white and Asian students to identify differences by race.

It is also worth noting that these results are not driven by differential attrition out of 9th grade near the threshold. The last row of Table 9 uses as an outcome an indicator for being observed in a WCPSS high school at all. Well over 90% of WCPSS 8th graders near the threshold appear in a WCPSS high school. There is no evidence from these coefficients that the threshold has any impact on this probability, for all students and for each subgroup.

6 Discussion and Conclusion

Recently, the Wake County Public School System implemented a criterion-based strategy for placing students in the advanced mathematics course sequence in the middle grades. Applying a regression discontinuity strategy to data on multiple cohorts of students in the Wake County Public School System, we examine the impact of the district's middle grades course placement policy. Because the EVAAS-based course placement recommendations were not followed with complete fidelity, we utilize an instrumental variables approach to handle the "fuzzy" discontinuity.

Our analyses reveal several key findings. First, being accelerated in mathematics did not have a clear impact on students' performance on end-of-grade standardized test performance in mathematics, although it did have a negative impact on students' course performance. Second, we importantly do not observe significant variation in the impact of acceleration across outcomes according to salient student characteristics such as race and socioeconomic status. Third, while acceleration did not impact course performance (as measured by course grades) among boys, it negatively impacted course performance among girls. While girls just below the EVAAS threshold outperform their male peers, course performance was similar among boys and girls with EVAAS scores just above the threshold. Finally, for those cohorts that we can observe progressing into high school, the impact of acceleration persists in terms of subsequent course placement. We consider each of these key findings in turn.

The lack of impact on standardized test performance may be driven by a variety of factors. First, we recognize that our estimates pertaining to standardized test performance are imprecisely estimated with large standard errors. Therefore, we are unable to rule out the possibility of modest impacts that we lack the precision to detect. From a substantive perspective, another possibility is that curricular differences between the advanced and non-advanced course sequences are modest, such that students are exposed to similar material in both levels and therefore are similarly prepared for the end-of-course assessments, regardless of course level. Yet a third possibility, however, is that the end-of-grade tests are not designed to be sensitive to the curricular differences that do exist between these two course levels.

Regarding the lack of impact variation by factors such as race and socioeconomic status, we highlight that the district's motivation for implementing the new policy was driven by goals of both access and equity. First, the district sought to increase the share of academically-ready students taking algebra by 8th grade so that these same students could access a college-preparatory mathematics curriculum in high school. Second, the district was particularly concerned about the underrepresentation of black students, Hispanic students and students from low-income backgrounds in advanced coursework, particularly given patterns that students in these demographic groups who were academically prepared for advanced mathematics coursework in the middle grades were not otherwise taking it up. Encouragingly, the policy appears to have moved the district towards these goals. The new assignment policy increased the share of district students who complete Algebra I prior to high school. In addition, since the policy's implementation, the relationship between middle grades course assignment and student characteristics such as socioeconomic status and race / ethnicity has diminished while the relationship between course placement and prior academic achievement has become stronger (Dougherty et al, 2014). Further, we observe that acceleration impacted student outcomes similarly regardless of race / ethnicity or socio-economic status. This set of findings together points to the promise of policies such as this for improving students' trajectory of rigorous coursetaking and subsequent access to higher-level mathematics in high school.

Where we do observe variation, however, is in the impact of acceleration by gender. Specifically, acceleration has a detrimental impact on the short-term course performance of girls. Here, we relate this finding to the literature on gender and participation and success in competitive academic environments and mathematics performance, in particular. Interestingly, much of the literature that we touch on here focuses on explanations for gender differences in the labor market. Bertrand (2010) notes that many of the most recent pieces of evidence are drawn from lab-based experimental settings and have yet to be validated by "demonstration of [their] economic significance in real markets" (1546). Our findings make a key contribution to informing this gap in the literature.

While gender gaps in mathematics have changed over time, such that boys only narrowly outperform girls, on average, differentials in achievement at the upper end of the performance distribution are vastly different, and these differences begin as early as first grade (Robinson and Lubienski, 2011). For example, in the top five percent of scores on the mathematics SAT, boys are overrepresented, with a two-to-one gender ratio (Ellison and Swanson, 2010). Ellison and Swanson (2010) also note dramatic gender differences in the geographic distribution of children in the U.S. who are highest performing in mathematics. While top-performing boys come from a variety of places around the US, the top-performing girls attend one of about 20 high schools in the country. As the authors note, "this suggests that almost all American girls with extreme mathematical ability are not developing their mathematical talents to the degree necessary to reach the extreme top percentiles in these contests" (110). Therefore, understanding the decisions and experiences of girls as they face the opportunity to participate in a higher-level mathematics trajectory is of critical importance.

Standard explanations for gender gaps in mathematics include factors such as low parental and educator expectations, less investment by girls in mathematics, biased testing, and unobserved ability. Data from a variety of countries (Bharadwag et al, 2012) and on representative data from the US (Fryer & Levitt, 2009) generally comport little with these hypothesized mechanisms. Therefore, more recently, economists have turned to psychological explanations for gender differences in mathematics performance.

First, research has documented gender differences in perceptions of math and one's own math ability. Gender gaps in math performance appear early in children's educational trajectories and appear to grow over time (Bharadwag et al, 2012; Robinson & Lubienski, 2011). Similarly, there are gender gaps in children's feelings towards math and perceptions of own ability in math, with girls having more negative math attitudes than boys, even after controlling for math performance (Ruble et al, 1993; Bharadwag et al, 2012; Gunderson et al 2012). These feelings are not necessarily related exclusively to math, however. When examining a more global measure of self-worth, for example, Kling and colleagues (1999) find that girls have a more negative self-image of themselves, and that this is particularly true in adolescence. Particularly in school settings, girls are more likely to experience anxiety. As with the math gender gaps, gaps in experienced anxiety and internal distress also grow over time (Lewinsohn et al, 1998). Related to the focus of this investigation on mathematics in the middle grades, girls may be at the highest risk for internal distress at times of challenge, such as when transitioning to a new school (Angold & Rutter, 1992; Simmons & Blyth, 1987) or when faced with an unfamiliar task (Ruble et al, 1993). Given that Wake County has an elementary, middle and high school structure, exposure to the higher-level mathematics trajectory coincides for students with the transition to a new school building, and the higher-level mathematics course may present students with a greater number of unfamiliar academic tasks or concepts. Therefore, the timeline of the policy may be coincident with increased risk for internal distress among girls. Similarly, the higher-level mathematics trajectory may be more likely to induce internal distress among female students, thus impacting their performance in the higher-level course.

Another body of literature that serves to shed light on how responses to the mathematics policy may differ by gender focuses on gendered responses to competitive environments. In a lab-based tournament experiment, Niederle and Vesterlund (2010) find that women are less likely to opt into a competitive setting, specifically when they can observe the gender make-up of potential competitors. Even conditional on performance measured prior to the opportunity to select the tournament setting, the authors find that men are more likely to select the competitive option and are therefore judged to be more overconfident in their actual skills (Niederle & Vesterlund, 2007). This finding corroborates psychological research on gender differences in overconfidence (Lundeberg et al, 1994; Beyer, 1990; Beyer & Bowden, 1997). Preckel and colleagues (2008) observe that gender gaps in academic confidence are greatest among the most gifted children, and Niederle and Vesterlund (2010) argue that "confidence and attitudes toward competition are likely to influence performance on competitive math tests and that these differences may play a substantial role at the right tail of the distribution" (137). This literature suggests that there may be particular differences in academic confidence when comparing boys and girls in the higher-level mathematics trajectory.

We observe a modest (and suggestive) difference in take up rates at the EVAAS margin. This may be because girls expected the higher-level math trajectory to have larger shares of male students or to be a more competitive academic environment. Related to confidence, girls may have been less likely to believe that their placement in the higher-level trajectory was accurate. For girls who took up the opportunity to be in the higher-level trajectory, if girls experienced the environment as more competitive, this could also have led to a suppression of their performance. In a lab-based experiment, Gneezy, Niederle and Rustichini (2003) find that in mixed-gender competitions, women fail to perform well, although the same is not true when they are competing in single-gender competitions or in settings where the gender of their competitors is unknown. Huguet and Regner (2007) similarly find that girls mathematics performance is sensitive to gender composition. The authors find that in tasks measuring mathematics ability, girls underperform in mixed-gender but not all-female groups. Further, they observe that attitudes towards tasks vary according to the gender composition of the group.

Also related to attitudes toward competition, Booth and Nolen (2009a, 2009b) find that after controlling for background characteristics and instrumenting for type of school attended, girls in mixed-gender schools are more risk-averse compared to their counterparts in single-gender schools. Fryer and Levitt (2009) find that in Muslim countries, where children are more likely to be educated in single-sex environments, there is little gender gap in mathematics. Finally, at the age at which the policy impacts students, girls may also be sensitive to gender identity and gender stereotyping. When situated in domains that are stereotypically male-dominated, such as a higher-level mathematics course, girls may feel more pressure to take on more stereotypical gender identities (Maccoby, 1990, 1998). These differentials in norms may also lead girls to be less likely to take up the offer to enroll in the advanced math trajectory.

Together, these results suggest that middle school math acceleration has promise for increasing college readiness, though girls' math performance may suffer in settings where they are near the bottom of the relative skill distribution.

References

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association 91*(434), 444–455.

Calonico, S., M. D. Cattaneo, and R. Titiunik (forthcoming). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review*, 158–165.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources 42*(3), 528–554.

Dougherty, S., J. Goodman, D. Hill, E. Litke, and L. Page (2014). Middle school math acceleration and equitable access to 8th grade algebra: Evidence from the Wake County public school system. Working Paper RWP14-029, Harvard Kennedy School.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies 79*(3), 933–959.

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics 142*(2), 615–635.

Ludwig, J. and D. L. Miller (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, 159–208.

Figure 1: Fraction of Students Accelerated, By Year and Eligibility

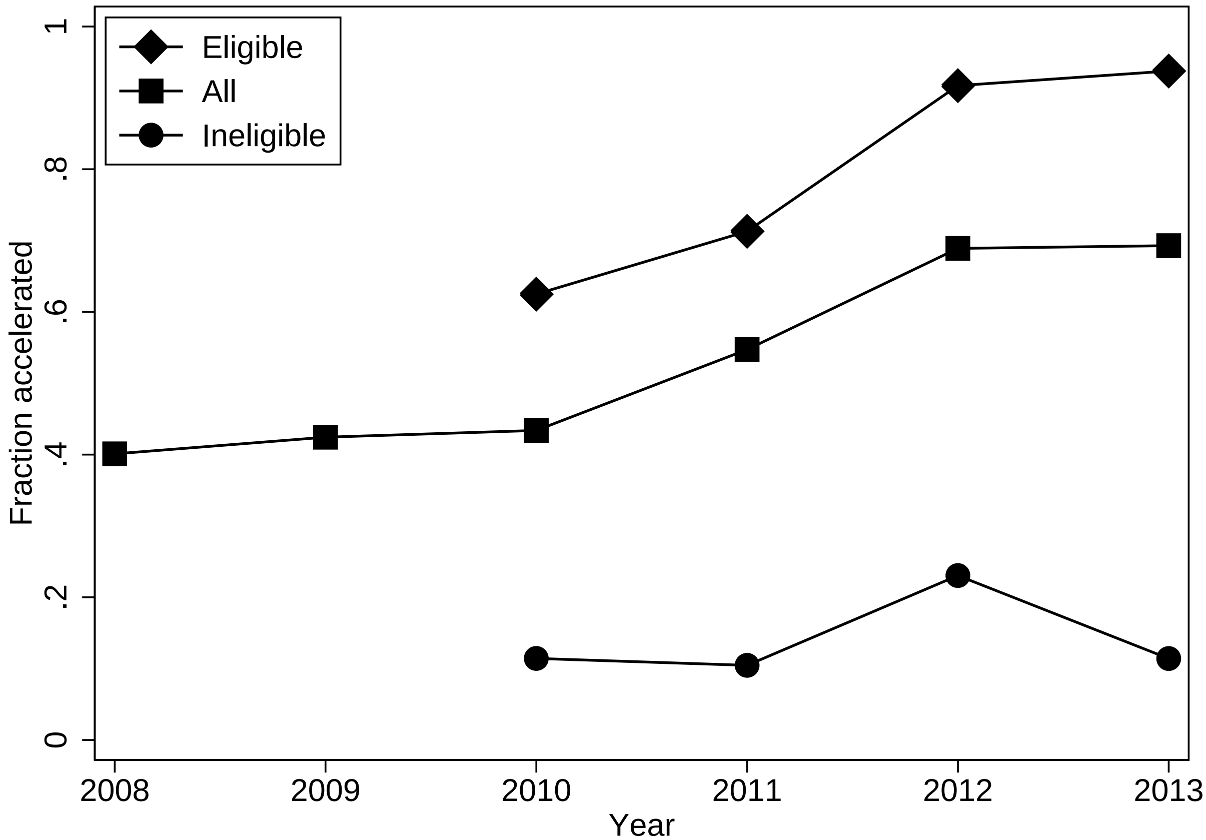


Figure 2: Distribution of EVAAS Scores

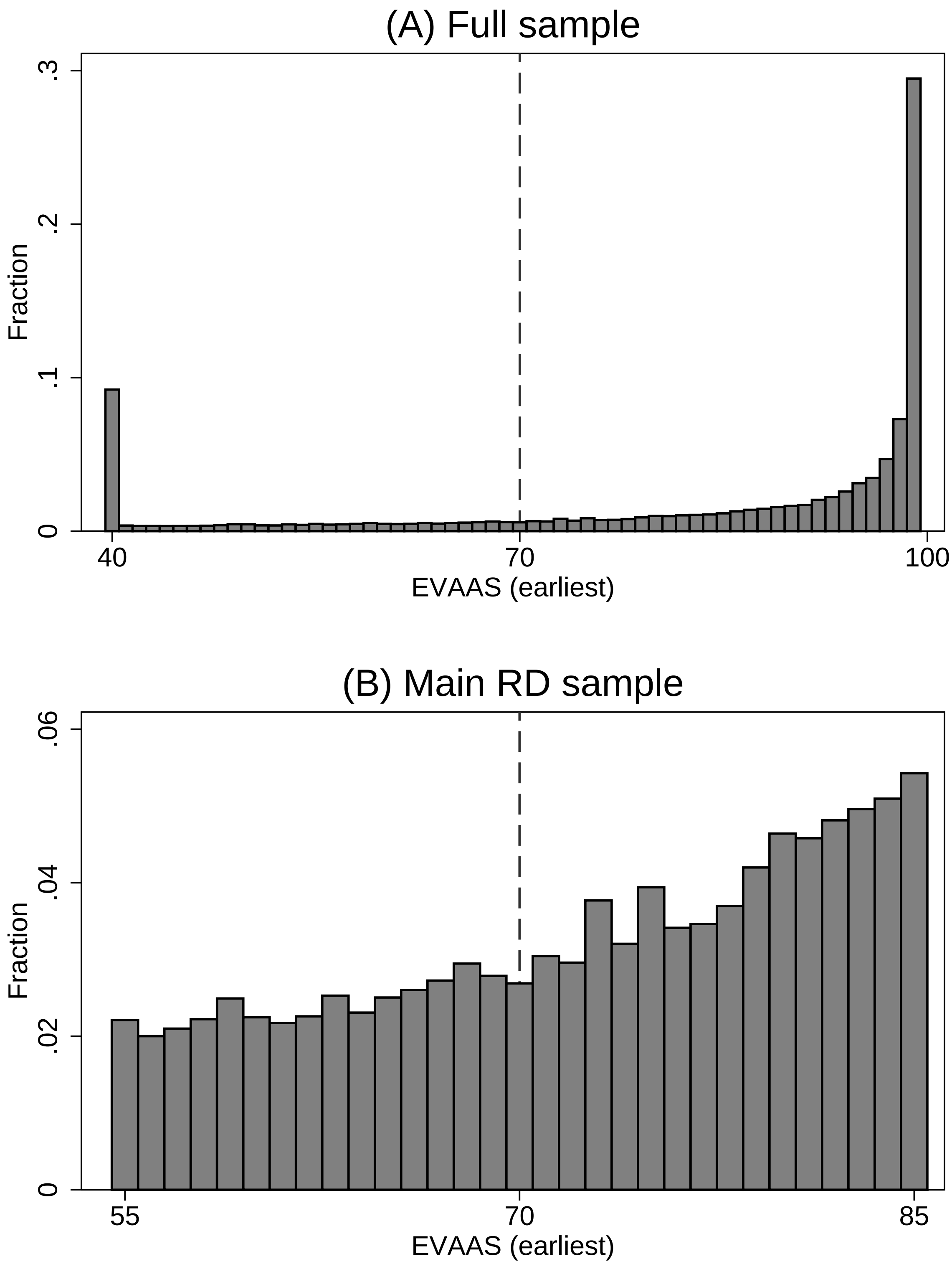


Figure 3: Placement in Accelerated Math, by Earliest EVAAS Score

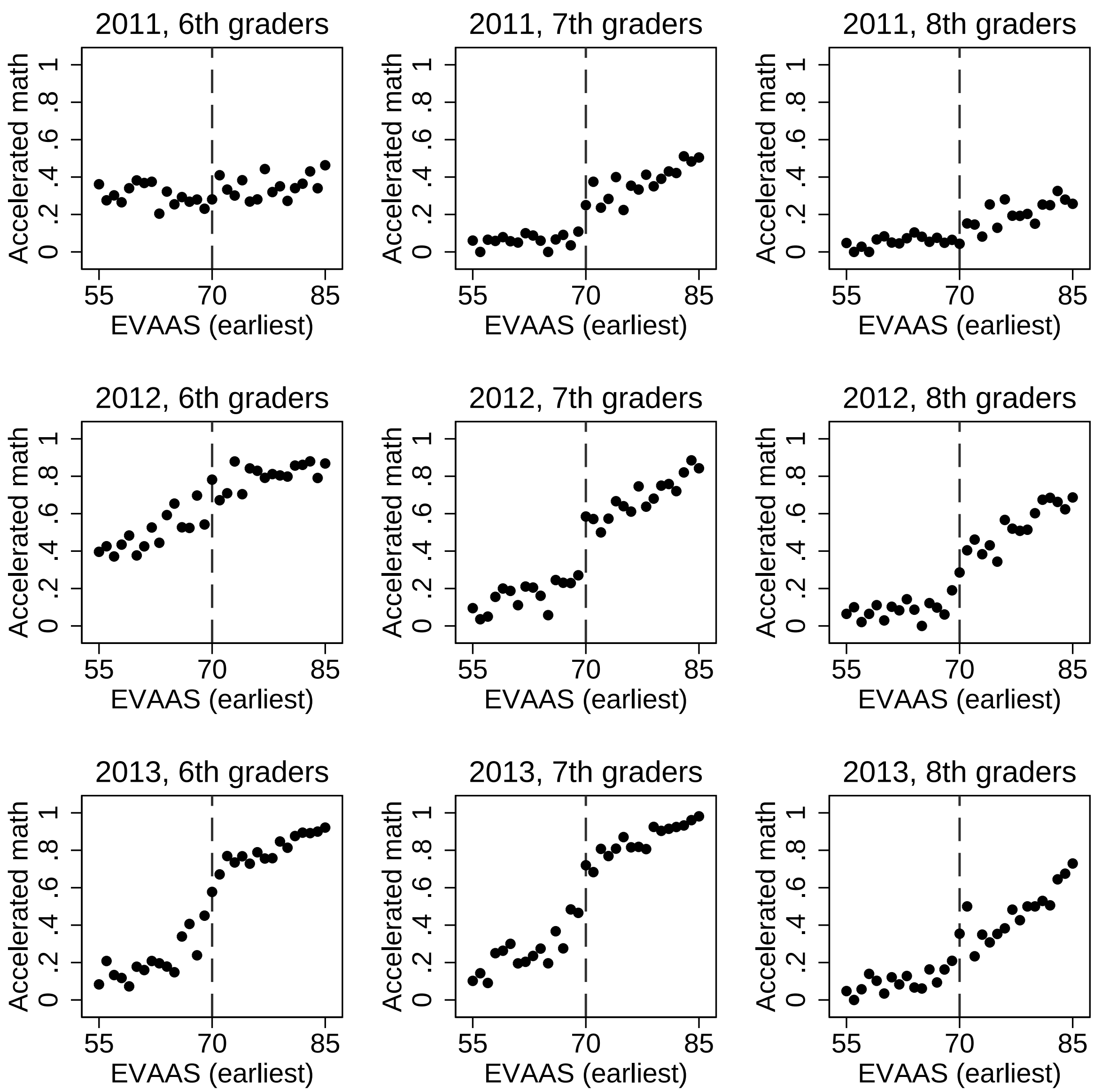


Figure 4: Fraction of Years in Accelerated Math, by Earliest EVAAS Score

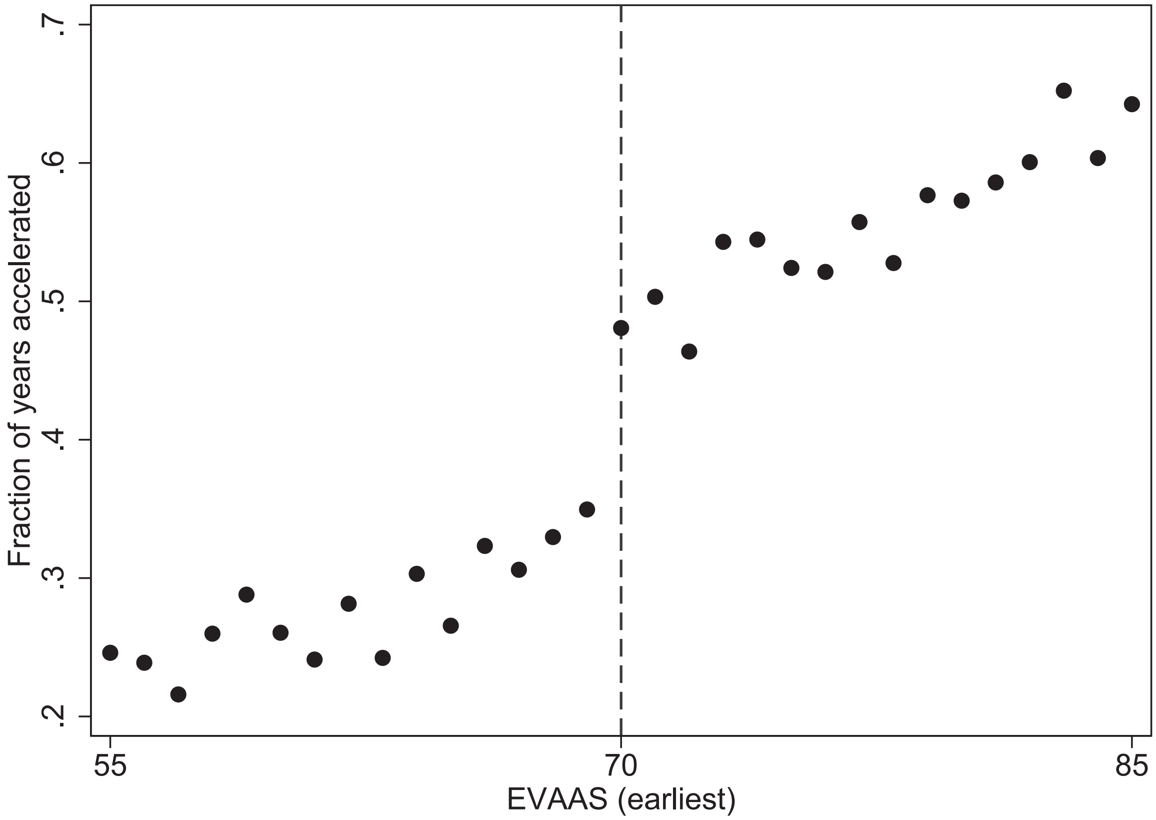


Figure 5: Impact of Acceleration on Peer and Teacher Characteristics

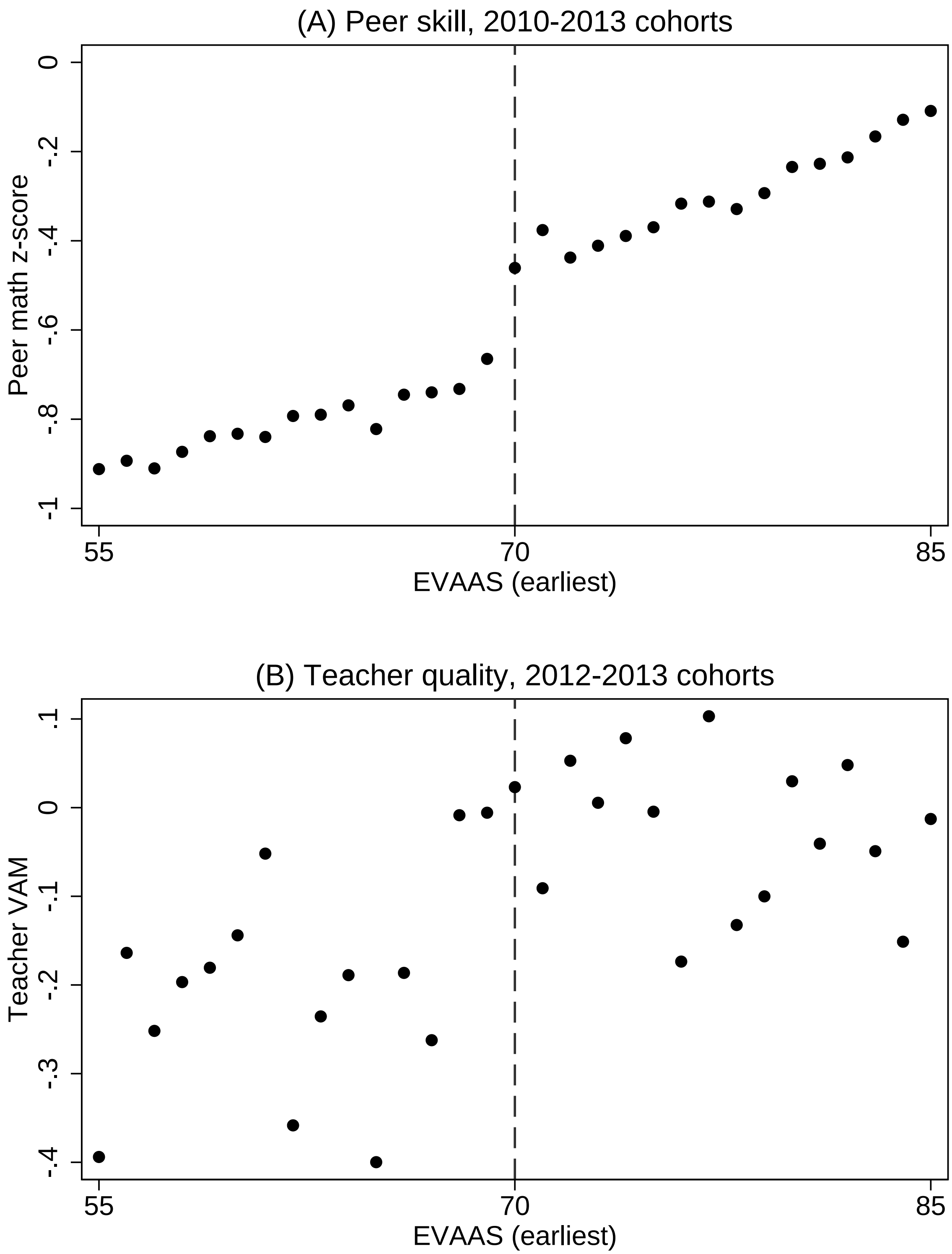


Figure 6: Math GPA, by Earliest EVAAS Score

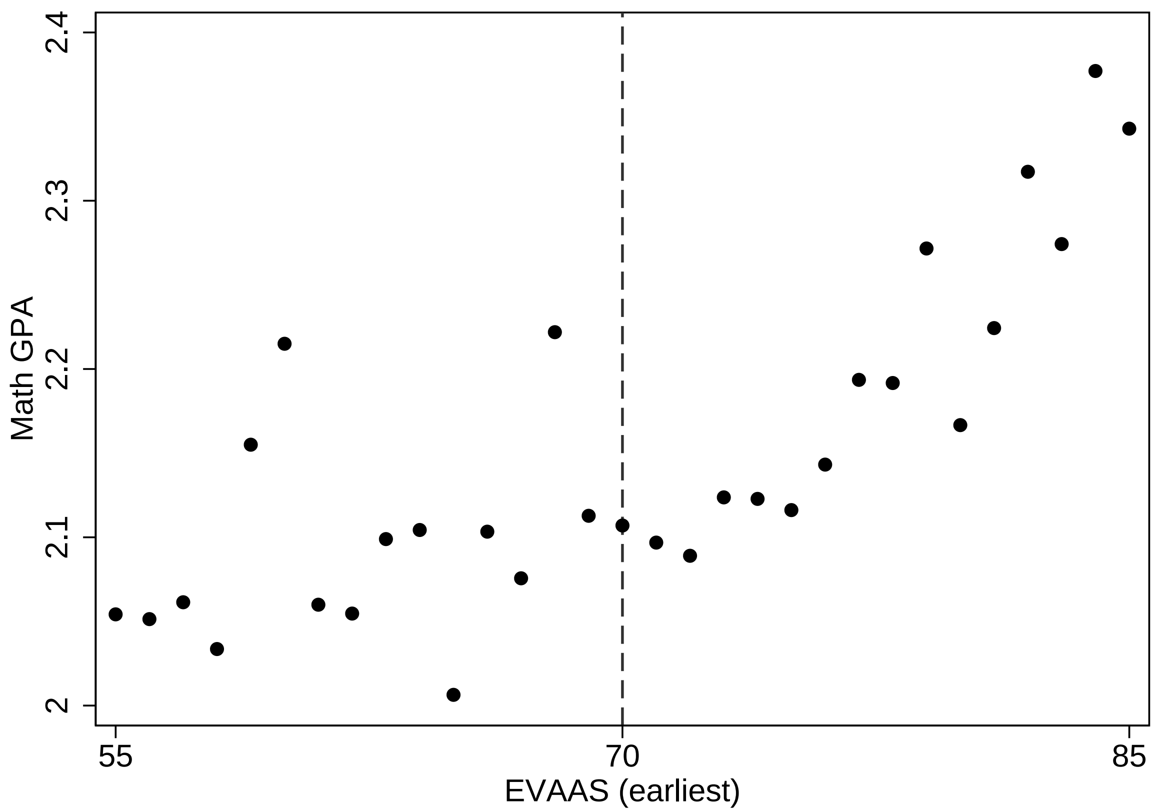


Figure 7: Math Z-Score, by Earliest EVAAS Score

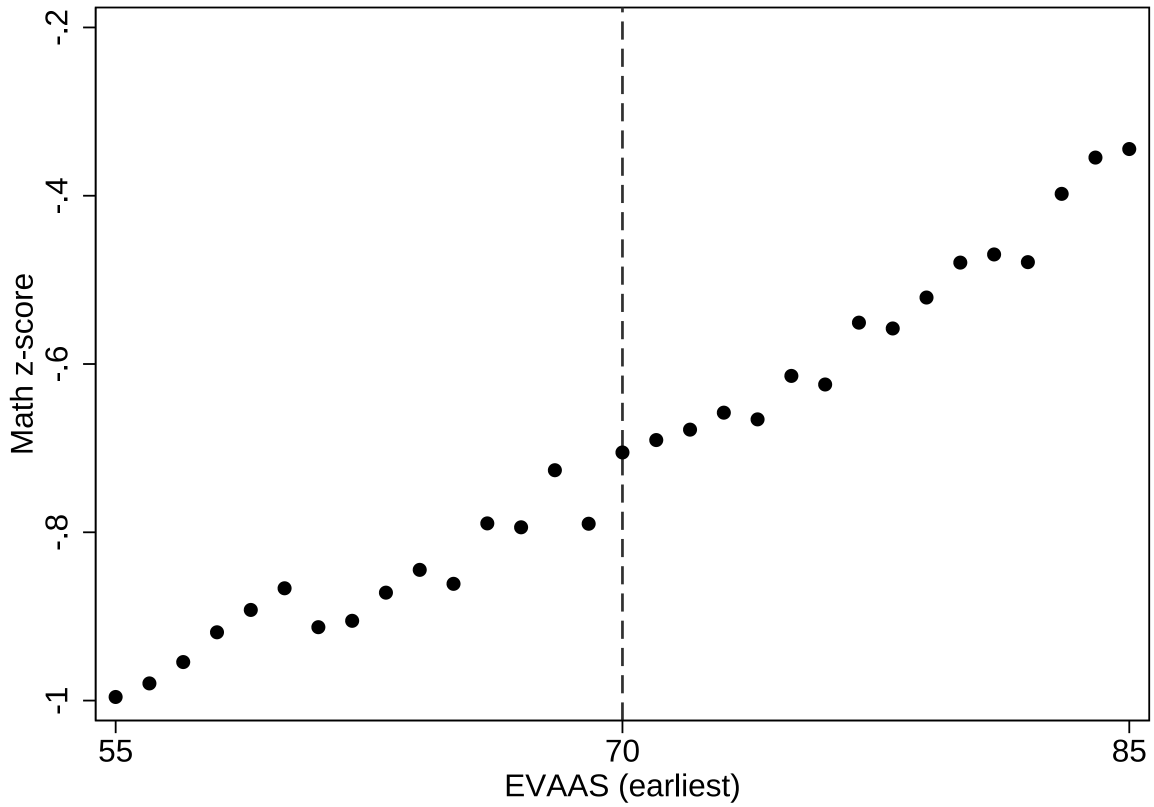


Figure 8: Math GPA, by Gender

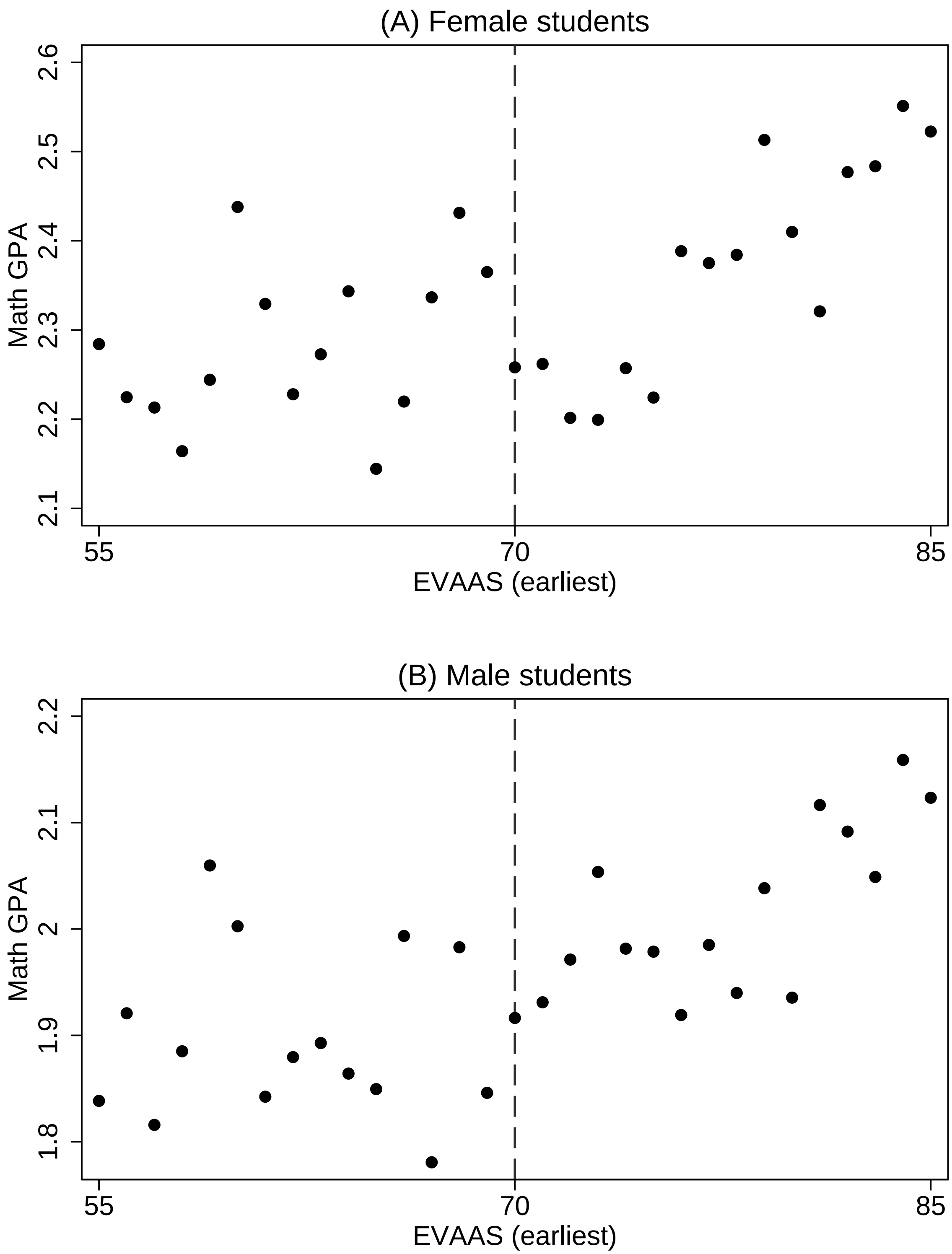


Figure 9: Math Z-Score, by Gender

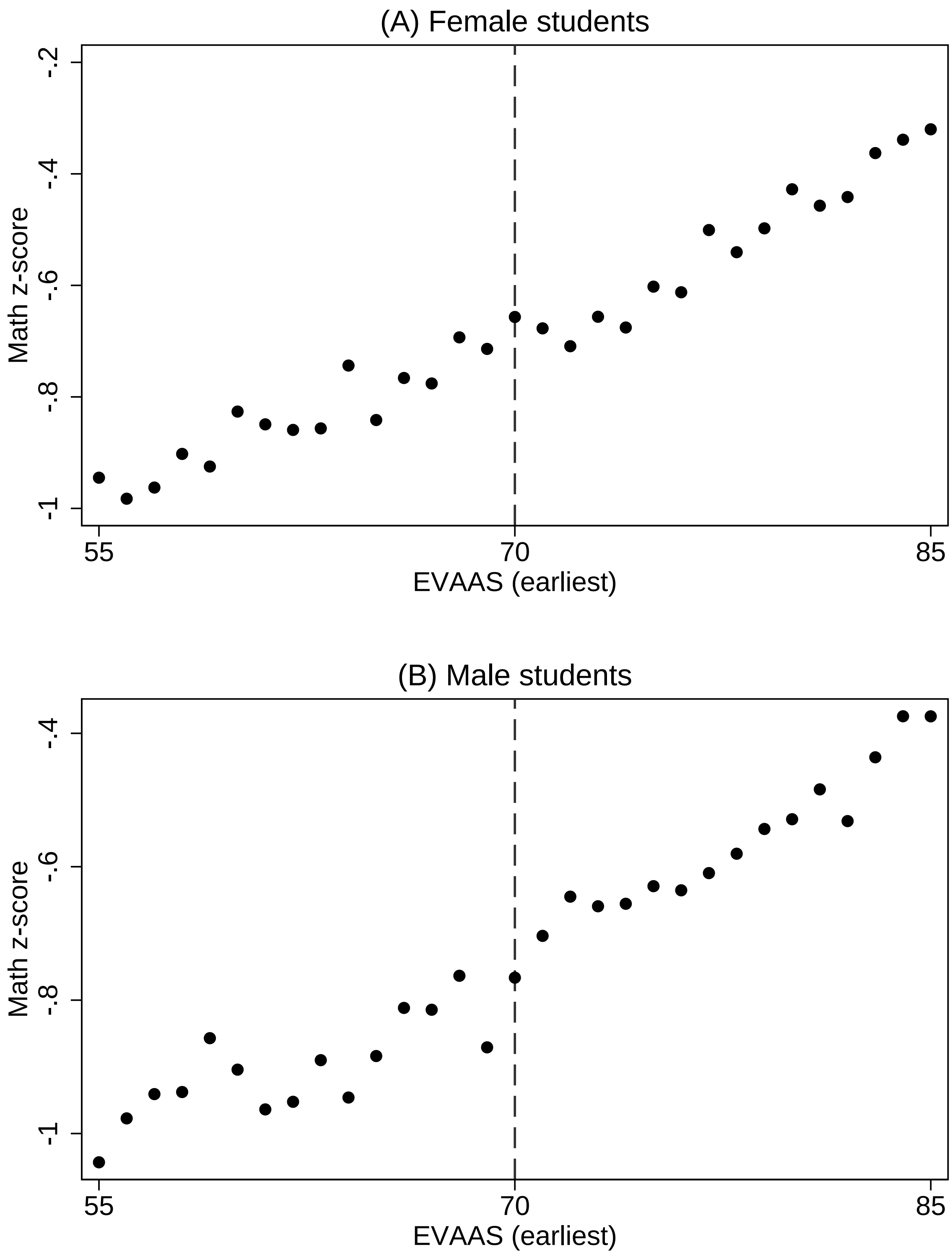


Table 1: Summary Statistics

	(1) All students	(2) Accelerated	(3) Non-accelerated
(A) Controls			
Female	0.501	0.504	0.495
White	0.509	0.598	0.304
Asian	0.062	0.079	0.024
Black	0.247	0.171	0.425
Hispanic	0.135	0.106	0.203
Other race	0.046	0.046	0.044
Poor	0.375	0.266	0.630
Special education	0.351	0.366	0.316
Limited English proficiency	0.161	0.138	0.213
Age on September 1	13.299	13.262	13.386
(B) Math course and skills			
Accelerated	0.699	1.000	0.000
EVAAS (most recent)	80.671	91.824	54.171
EVAAS (earliest)	83.051	92.868	60.235
5th grade math z-score	0.032	0.421	-0.884
(C) Math course peer composition			
Mean 5th grade math z-score	0.018	0.412	-0.898
SD 5th grade math z-score	0.629	0.617	0.655
Class size	26.243	27.719	22.811
Fraction black or Hispanic	0.429	0.327	0.666
Fraction female	0.499	0.502	0.492
(D) Grade and test score outcomes			
Math GPA	2.726	2.973	2.154
Passed math class	0.960	0.979	0.916
At least B in math class	0.626	0.722	0.403
End-of-grade math z-score	0.068	0.423	-0.790
N	82,359	57,584	24,775

Notes: Mean values of key variables are shown for all students in the 2010-2013 cohorts.

Table 2: McCrary and Covariate Balance Test

	(1) Count	(2) Black/ Hispanic	(3) Low income	(4) Special ed.	(5) LEP	(6) Age on Sept. 1	(7) Female
(A) All students							
Eligible	-2.573	0.000	-0.034	-0.032	0.020	0.017	-0.002
	(2.600)	(0.025)	(0.028)	(0.025)	(0.018)	(0.031)	(0.027)
μ	52.52	0.70	0.65	0.22	0.21	13.30	0.53
N	261	16,010	16,010	16,010	16,010	16,010	16,010
(B) Males							
Eligible	-0.796	-0.002	-0.009	-0.028	0.002	0.020	
	(1.554)	(0.036)	(0.044)	(0.038)	(0.034)	(0.046)	
μ	24.89	0.72	0.65	0.26	0.23	13.36	
N	261	7,677	7,677	7,677	7,677	7,677	
(C) Females							
Eligible	-1.712	-0.003	-0.065*	-0.033	0.045**	0.018	
	(1.743)	(0.033)	(0.037)	(0.030)	(0.019)	(0.035)	
μ	27.63	0.69	0.65	0.19	0.20	13.25	
N	261	8,333	8,333	8,333	8,333	8,333	

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each coefficient is the reduced form estimate of the relationship between eligibility for acceleration and the list covariate. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. The first column uses as an outcome the number of observations in each year, grade and 1-point wide EVAAS bin. Panel A includes all students, while panels B and C split the sample by gender. Also listed is the mean of the covariate for students just below the threshold (with EVAAS between 67 and 70). The sample includes the 2010-2013 cohorts.

Table 3: First Stage Estimates of Exposure to Accelerated Math Coursework

	(1) 2009 cohort	(2) 2010 cohort	(3) 2011 cohort	(4) 2012 cohort	(5) 2013 cohort	(6) 2010-13 cohorts
Fractions of years accelerated	0.012	0.073***	0.092***	0.186***	0.283***	0.134***
	(0.022)	(0.024)	(0.029)	(0.034)	(0.072)	(0.019)
μ	0.21	0.20	0.25	0.55	0.37	0.33
F	0.3	9.1	10.2	29.6	15.5	48.9
N	4,736	4,910	4,956	4,010	2,134	16,010
Accelerated in grade 6	0.016	0.004	0.012	0.158***	0.283***	0.125***
	(0.033)	(0.023)	(0.025)	(0.053)	(0.072)	(0.027)
μ	0.31	0.28	0.26	0.59	0.37	0.39
F	0.2	0.0	0.2	9.0	15.5	21.2
N	1,554	1,679	1,763	2,098	2,134	7,674
Accelerated in grade 7	-0.005	0.189***	0.276***	0.291***		0.252***
	(0.023)	(0.069)	(0.065)	(0.050)		(0.039)
μ	0.06	0.08	0.24	0.42		0.25
F	0.0	7.6	17.7	34.3		41.6
N	1,594	1,669	1,676	1,912		5,257
Accelerated in grade 8	0.049*	0.189***	0.100*			0.147***
	(0.028)	(0.058)	(0.055)			(0.035)
μ	0.06	0.11	0.16			0.14
F	3.2	10.8	3.2			18.0
N	1,588	1,562	1,517			3,079

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). In the first row, first stage estimates show the impact of eligibility for acceleration on the fraction of middle school years spent in accelerated math coursework. The remaining rows show the impact of eligibility for acceleration on current acceleration status, by grade. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. Eligibility is measured by each student's earliest EVAAS score. Below each coefficient is the proportion of students just below the threshold enrolled in accelerated coursework, as well as the F-statistic associated with the excluded instrument.

Table 4: Peer and Teacher Characteristics in Primary Math Classroom

	(1)	(2)	(3)	(4)	(5)	(6)
(A) Peers	Mean math skill	St. dev. math skill	Class size	Fraction female	Fraction low inc.	Fraction black/Hisp.
Accelerated	1.079***	-0.016	4.154***	0.007	-0.241***	-0.225***
	(0.088)	(0.037)	(1.097)	(0.031)	(0.026)	(0.031)
N	16,010	16,010	16,010	16,010	16,010	16,010
(B) Teachers	VAM estimate	Low VAM	Years of exp.	Novice teacher	Female teacher	Missing teacher
Accelerated	0.284	-0.213**	0.138	0.014	0.142	-0.038
	(0.298)	(0.100)	(1.223)	(0.040)	(0.086)	(0.056)
N	14,110	14,110	12,649	12,649	12,713	16,010

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Each panel shows instrumental variables estimates of the impact of acceleration, where acceleration is instrumented with eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. In panel B, low VAM is defined as having an estimated VAM more than one standard deviation below average, and the final column's outcome is an indicator for missing information about a student's primary math teacher. The sample includes the 2010-13 cohorts.

Table 5: Course Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Math GPA				All grades	
	All grades	6th grade	7th grade	8th grade	Passed	At least B	Non-math GPA
2010-13 cohorts	-0.541**	-0.554	-0.826***	0.096	-0.108*	-0.289**	0.340
	(0.274)	(0.404)	(0.294)	(0.576)	(0.060)	(0.138)	(0.299)
μ	2.14	2.24	2.09	1.96	0.94	0.39	2.31
N	16,010	7,674	5,257	3,079	16,010	16,010	16,009
2010 cohort	0.302	29.945	-0.490	-0.083	-0.029	-0.204	1.116
	(0.884)	(178.285)	(0.802)	(0.999)	(0.210)	(0.454)	(1.057)
μ	2.13	2.15	2.25	1.99	0.94	0.38	2.31
N	4,910	1,679	1,669	1,562	4,910	4,910	4,910
2011 cohort	0.860	14.290	0.091	0.330	-0.131	0.423	2.116***
	(0.766)	(29.835)	(0.643)	(0.880)	(0.201)	(0.299)	(0.760)
μ	1.99	2.11	1.92	1.92	0.92	0.34	2.01
N	4,956	1,763	1,676	1,517	4,956	4,956	4,955
2012 cohort	-1.365***	-1.176**	-1.515***		-0.115	-0.558***	-0.837*
	(0.430)	(0.529)	(0.450)		(0.096)	(0.190)	(0.435)
μ	2.23	2.37	2.08		0.95	0.42	2.53
N	4,010	2,098	1,912		4,010	4,010	4,010
2013 cohort	-1.061***	-1.061***			-0.123	-0.542***	-0.003
	(0.402)	(0.402)			(0.077)	(0.188)	(0.205)
μ	2.29	2.29			0.96	0.44	2.54
N	2,134	2,134			2,134	2,134	2,134

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). Instrumental variables estimates show the impact of the fraction of middle school years spent in accelerated math courses on various measures of course grades, where that fraction is instrumented by eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. The outcome in columns 1-4 is GPA in math courses, measured on a 4-point scale. The outcomes in columns 5 and 6 are indicators for respectively passing and earning at least a B in math. The outcome in column 7 is GPA in non-math courses, measured on a 4-point scale. Below each coefficient is the mean of the outcome variable among students just below the threshold.

Table 6: Standardized Test Scores

	(1)	(2)	(3)	(4)	(5)	(6)
		Math z-scores			Took	Reading
	All grades	6th grade	7th grade	8th grade	math exam	z-score
2010-13 cohorts	0.142	-0.038	0.190	0.500*	-0.011	-0.219
	(0.119)	(0.156)	(0.178)	(0.299)	(0.032)	(0.244)
μ	-0.77	-0.74	-0.79	-0.80	0.99	-0.64
N	15,850	7,609	5,203	3,038	16,010	15,776
2010 cohort	0.621	6.473	0.849*	0.228	0.043	-1.043
	(0.449)	(38.508)	(0.490)	(0.471)	(0.096)	(0.977)
μ	-0.80	-0.81	-0.80	-0.79	0.99	-0.58
N	4,880	1,677	1,654	1,549	4,910	4,865
2011 cohort	0.321	0.857	-0.052	0.744*	0.068	0.711
	(0.389)	(2.860)	(0.391)	(0.397)	(0.109)	(0.757)
μ	-0.83	-0.84	-0.82	-0.82	0.98	-0.85
N	4,894	1,747	1,658	1,489	4,956	4,862
2012 cohort	-0.153	-0.436	0.066		-0.055	-0.513
	(0.177)	(0.295)	(0.234)		(0.050)	(0.357)
μ	-0.69	-0.62	-0.75		0.99	-0.51
N	3,963	2,072	1,891		4,010	3,952
2013 cohort	0.059	0.059			-0.047**	-0.067
	(0.134)	(0.134)			(0.023)	(0.270)
μ	-0.73	-0.73			1.00	-0.58
N	2,113	2,113			2,134	2,097

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). Instrumental variables estimates show the impact of the fraction of middle school years spent in accelerated math courses on standardized test scores, where that fraction is instrumented by eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. The outcome in columns 1-4 is an end-of-grade math exam z-score. The outcome in column 5 is an indicator for taking the math exam. The outcome in column 6 is an end-of-grade reading exam z-score. Below each coefficient is the mean of the outcome variable among students just below the threshold.

Table 7: Heterogeneity of Grade and Test Score Impacts, By Gender

	(1) Math GPA	(2) At least D	(3) At least B	(4) Math z-score
Male * Accelerated	0.060	-0.058	-0.043	0.359*
	(0.341)	(0.089)	(0.167)	(0.180)
Female * Accelerated	-1.364***	-0.177*	-0.623***	-0.157
	(0.369)	(0.098)	(0.199)	(0.193)
Female	0.971***	0.110	0.397***	0.263**
	(0.221)	(0.066)	(0.108)	(0.123)
μ (Male)	1.87	0.90	0.29	-0.82
μ (Female)	2.38	0.97	0.48	-0.73
p	0.01	0.44	0.02	0.08
N	16,010	16,010	16,010	15,850

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). Instrumental variables estimates show the impact of the fraction of middle school years spent in accelerated math courses on various measures of course grades and test scores, where that fraction is instrumented by eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. That specification is then interacted with indicators for each gender. Below each coefficient are the subgroup means of the outcome variable among students just below the threshold. Also shown is a p-value from an F-test of the equality of the two interaction coefficients. The sample includes the 2010-13 cohorts.

Table 8: Heterogeneity of Peer and Teacher Impacts, By Gender

	(1) Mean peer math skill	(2) Fraction female	(3) Fraction low income	(4) VAM estimate	(5) Female teacher
Male * Accelerated	1.339***	0.040	-0.281***	0.486	0.034
	(0.125)	(0.046)	(0.050)	(0.327)	(0.152)
Female * Accelerated	1.407***	-0.036	-0.341***	0.192	0.390*
	(0.136)	(0.057)	(0.064)	(0.835)	(0.202)
Female	0.004	0.079**	0.018	0.052	-0.146
	(0.071)	(0.030)	(0.040)	(0.379)	(0.121)
μ (Male)	-0.70	0.48	0.59	-0.12	0.79
μ (Female)	-0.63	0.54	0.58	0.06	0.76
p	0.69	0.28	0.53	0.75	0.22
N	16,010	16,010	16,010	14,110	12,713

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). Instrumental variables estimates show the impact of the fraction of middle school years spent in accelerated math courses on various measures of peer and teacher characteristics in the primary math classroom, where that fraction is instrumented by eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. That specification is then interacted with indicators for gender. Below each coefficient are the subgroup means of the outcome variable among students just below the threshold. Also shown is a p-value from an F-test of the equality of the two interaction coefficients. The sample includes the 2010-13 cohorts.

Table 9: High School Freshman Math Coursework

	(1) All students	(2) Female students	(3) Male students	(4) Low income	(5) Non-low income	(6) Black/ Hispanic	(7) White/ Asian
Enrolled in geometry	0.810***	0.921***	0.623***	0.732**	0.773***	0.854***	0.466
	(0.119)	(0.162)	(0.174)	(0.291)	(0.153)	(0.128)	(0.695)
μ	0.12	0.13	0.12	0.11	0.16	0.11	0.17
N	2,908	1,517	1,391	1,834	1,074	2,017	891
Passed geometry	0.717***	0.805***	0.483**	0.462**	0.848***	0.766***	0.433
	(0.150)	(0.237)	(0.207)	(0.234)	(0.201)	(0.163)	(0.766)
μ	0.09	0.09	0.10	0.08	0.12	0.07	0.16
N	2,908	1,517	1,391	1,834	1,074	2,017	891
A or B in geometry	0.064	0.150	0.009	-0.059	0.094	0.006	0.306
	(0.112)	(0.151)	(0.136)	(0.146)	(0.179)	(0.089)	(0.717)
μ	0.01	0.01	0.02	0.01	0.03	0.01	0.03
N	2,908	1,517	1,391	1,834	1,074	2,017	891
F	14.9	7.2	14.8	4.1	7.6	18.8	0.6
In WCPSS high school	0.028	0.019	0.049	0.015	0.014	-0.033	0.314
	(0.151)	(0.247)	(0.166)	(0.257)	(0.168)	(0.138)	(0.610)
μ	0.94	0.93	0.94	0.93	0.95	0.94	0.93
N	3,079	1,602	1,477	1,946	1,133	2,125	954

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). Instrumental variables estimates show the impact of acceleration in 8th grade on various measures of high school freshman math course enrollment, where acceleration is instrumented by eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. The sample consists of those members of the 2010 and 2011 cohorts who are present in the data in 8th grade. The first three rows use as outcomes indicators for enrolling in geometry (or a higher course), passing that course, and earning an A or B in that course. The final row indicates whether a student appears in the 9th grade data. Below each coefficient is the mean of the outcome variable among students just below the threshold. Also shown is a first stage F-statistic from a test for weak identification of the instrument.

Figure A.1: Fraction of Students Accelerated, By Gender

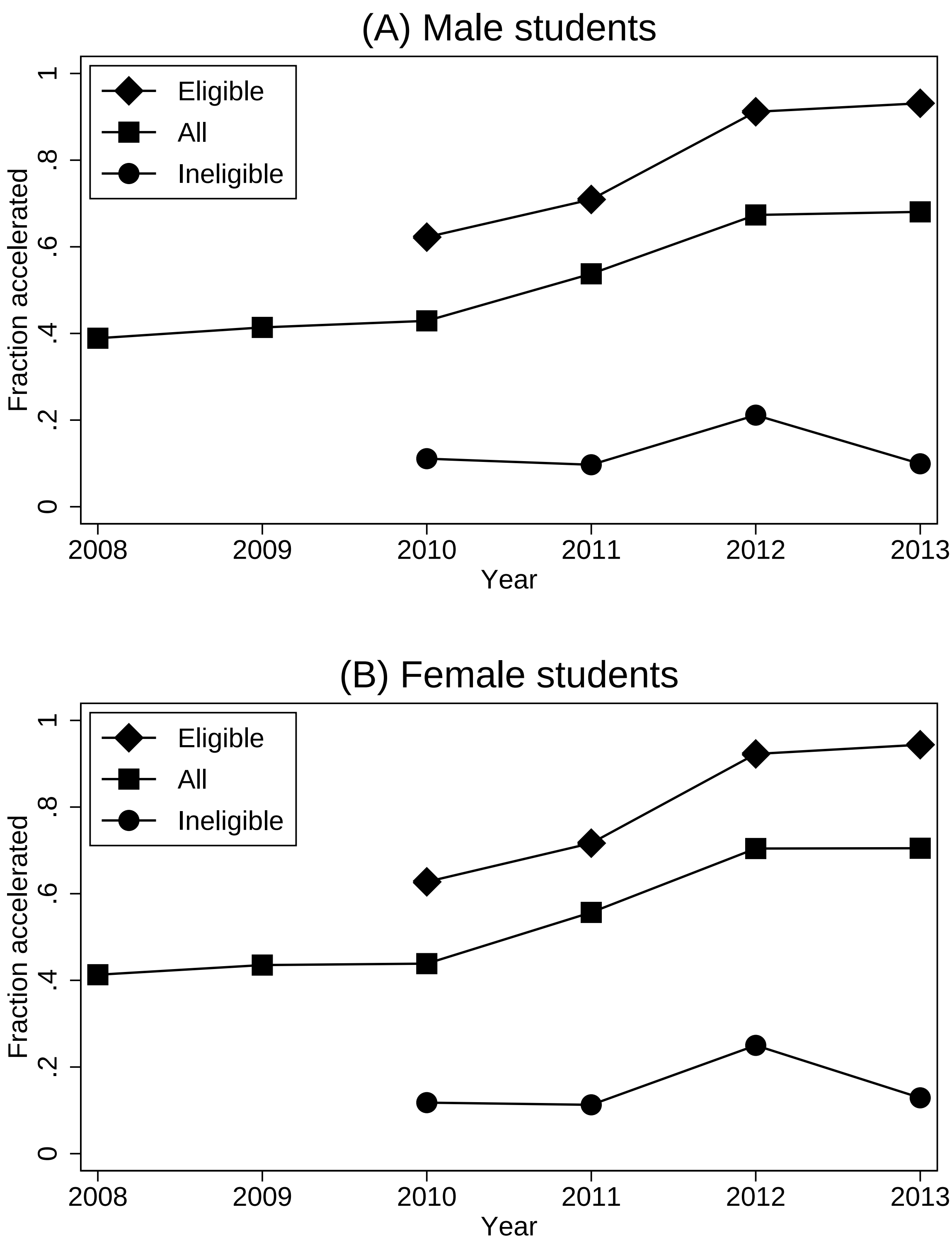


Figure A.2: Fraction of Students Accelerated, By Income

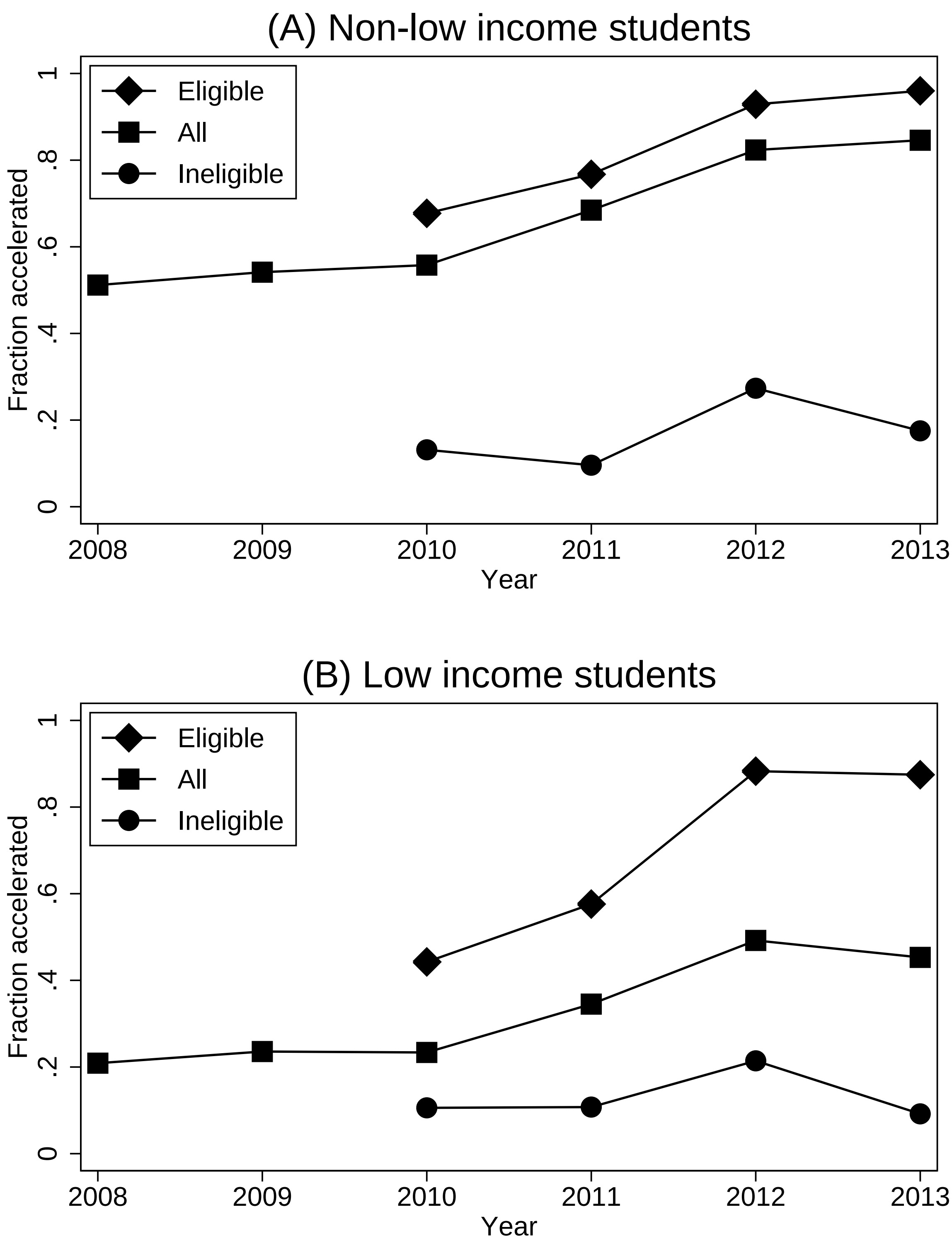


Figure A.3: Fraction of Students Accelerated, By Race

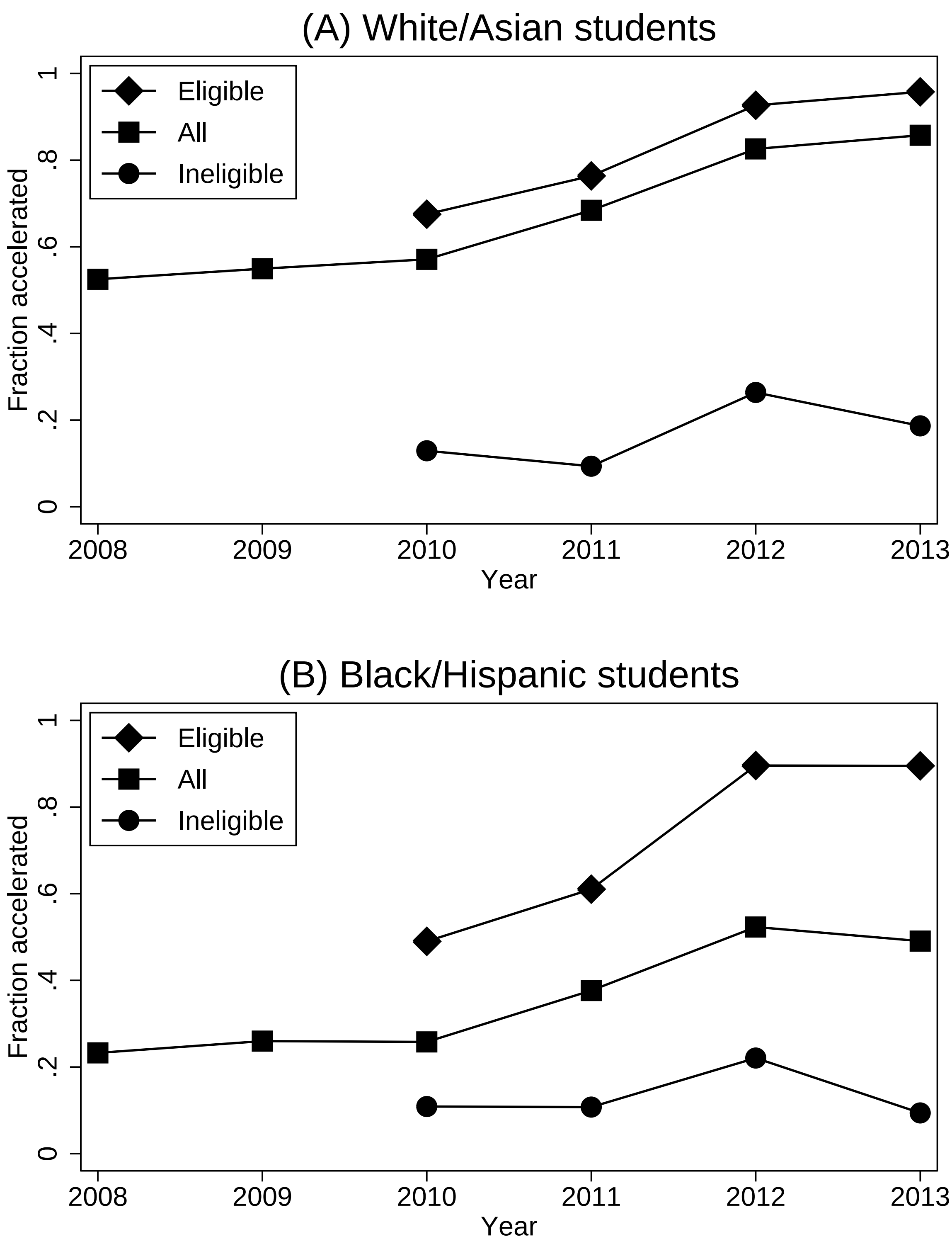


Figure A.4: Placement in Accelerated Math, by Gender

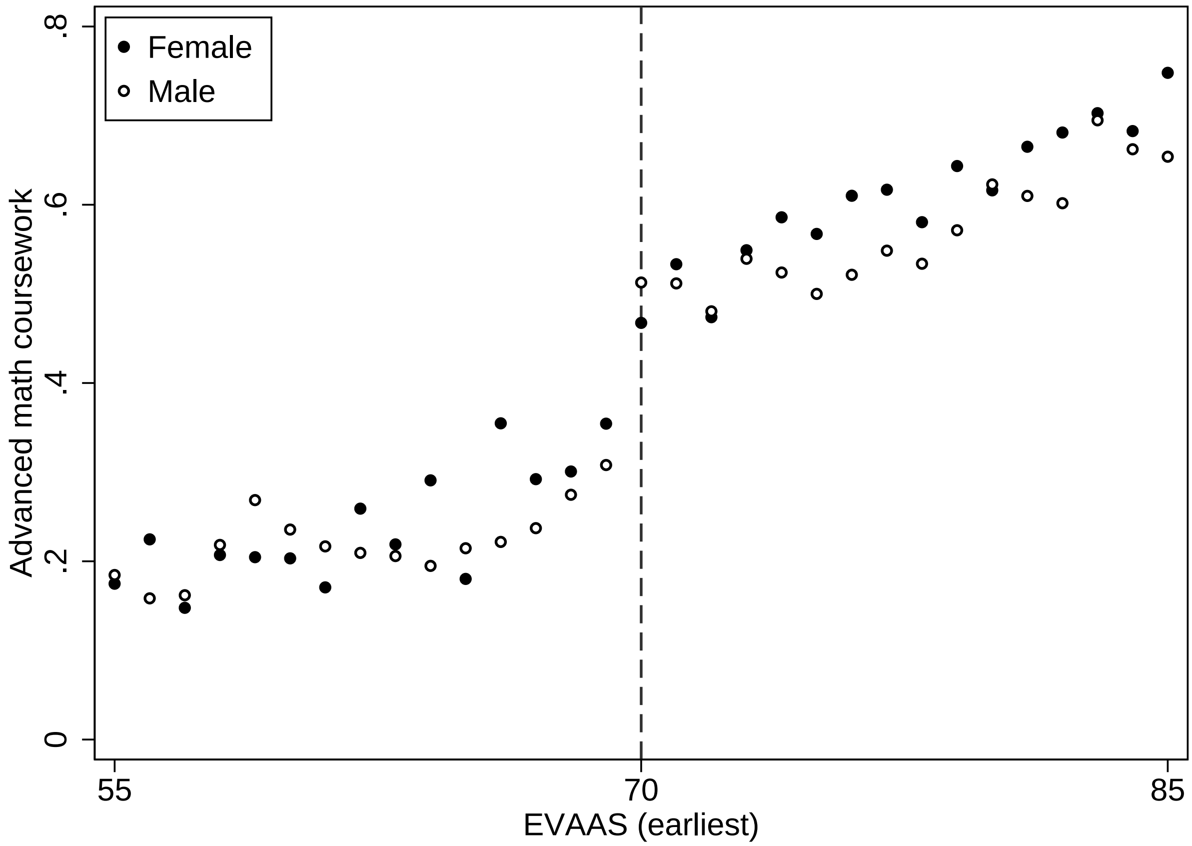


Figure A.5: Placement in Accelerated Math, by Income

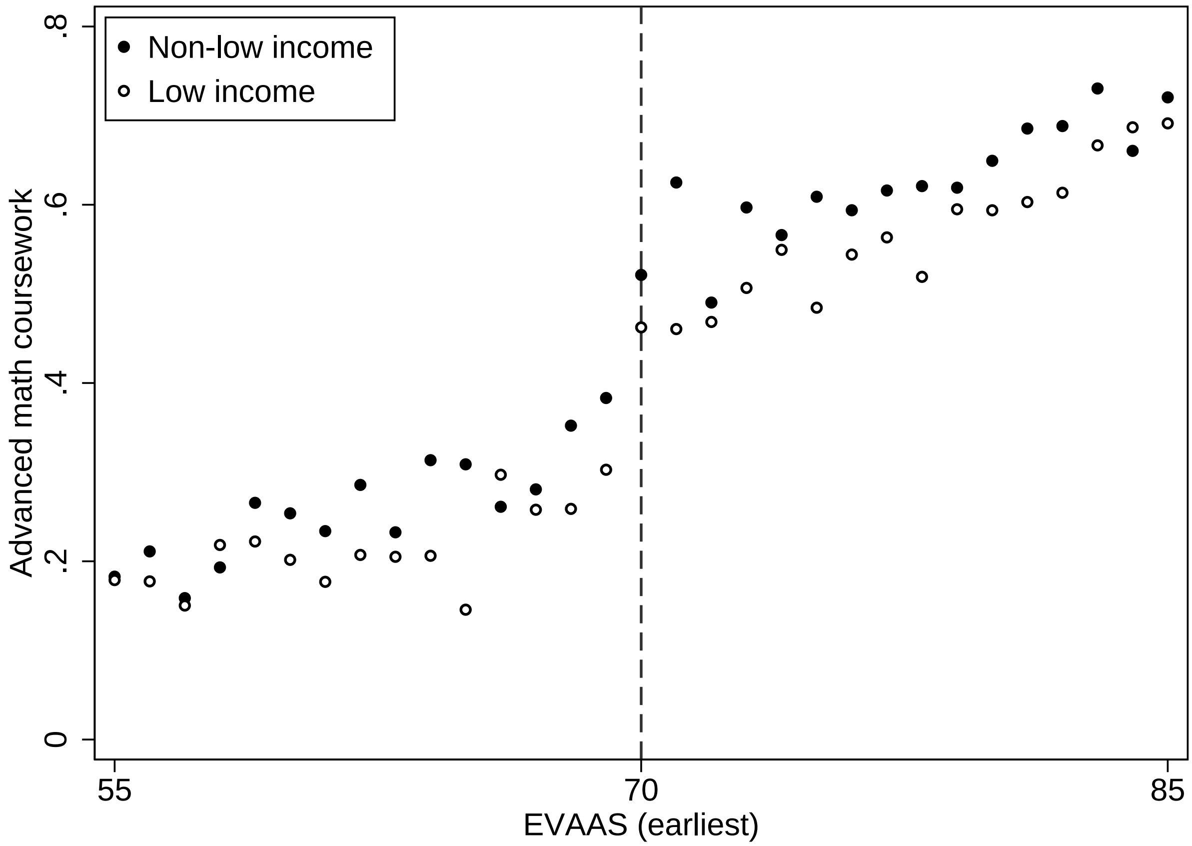


Figure A.6: Placement in Accelerated Math, by Race

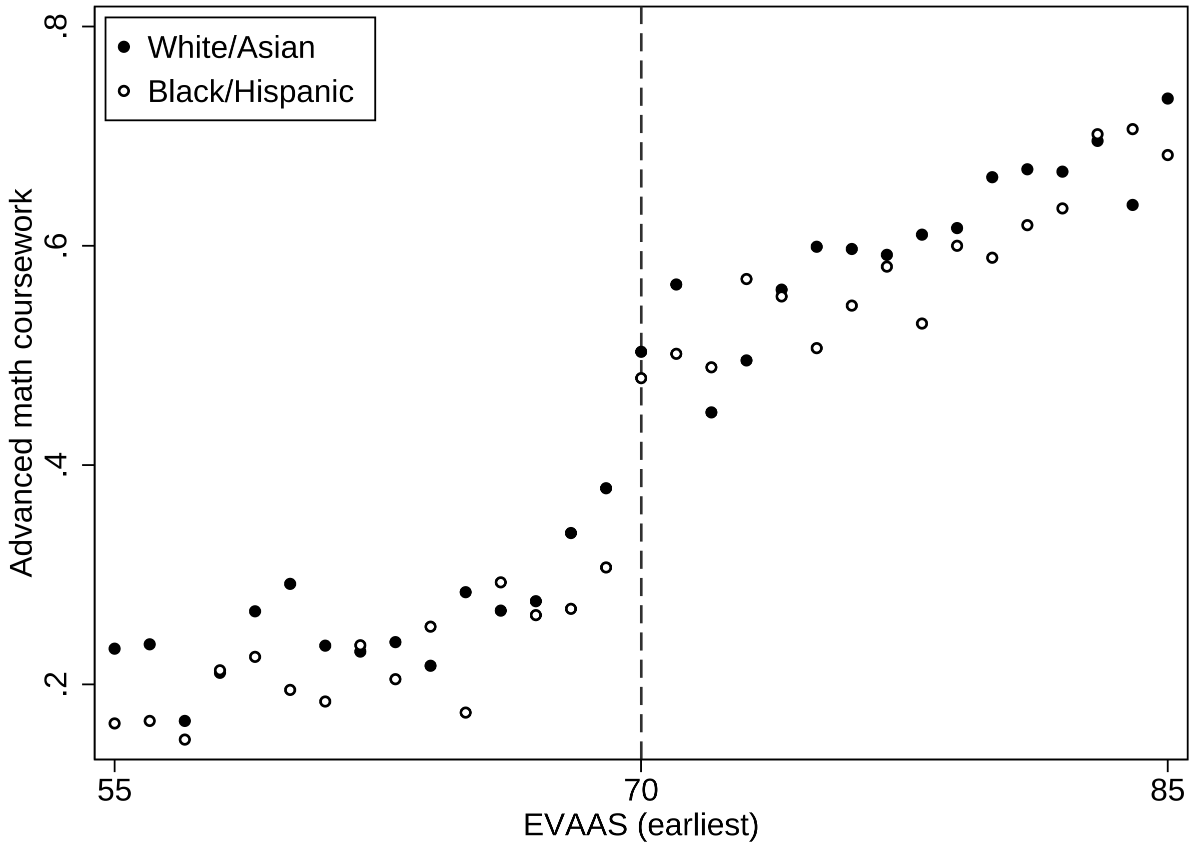


Figure A.7: Math GPA, by Income

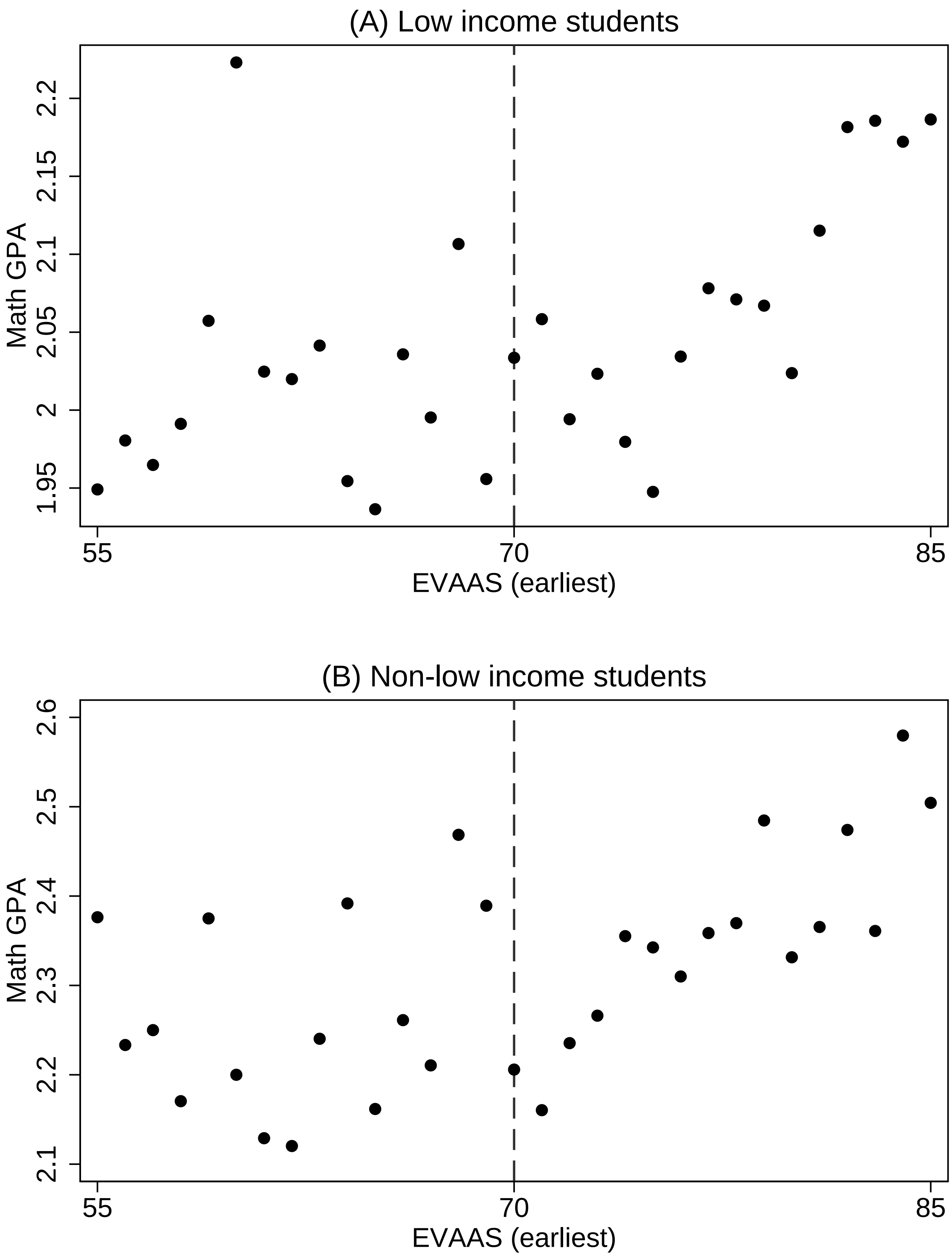
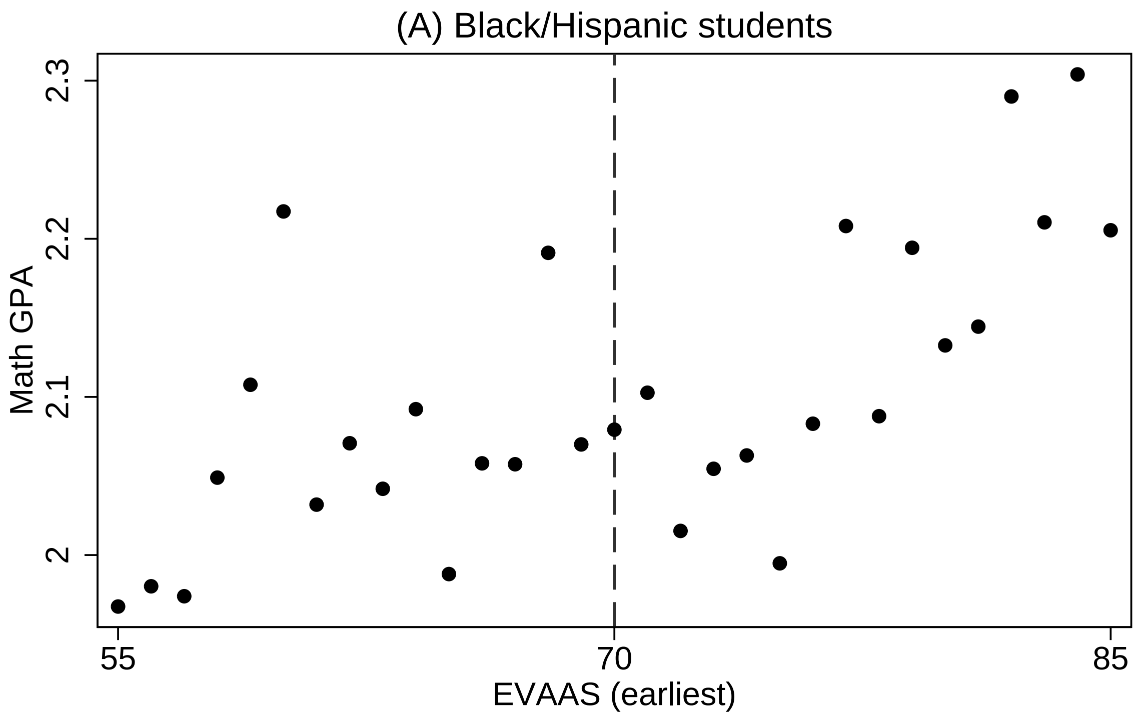


Figure A.8: Math GPA, by Race



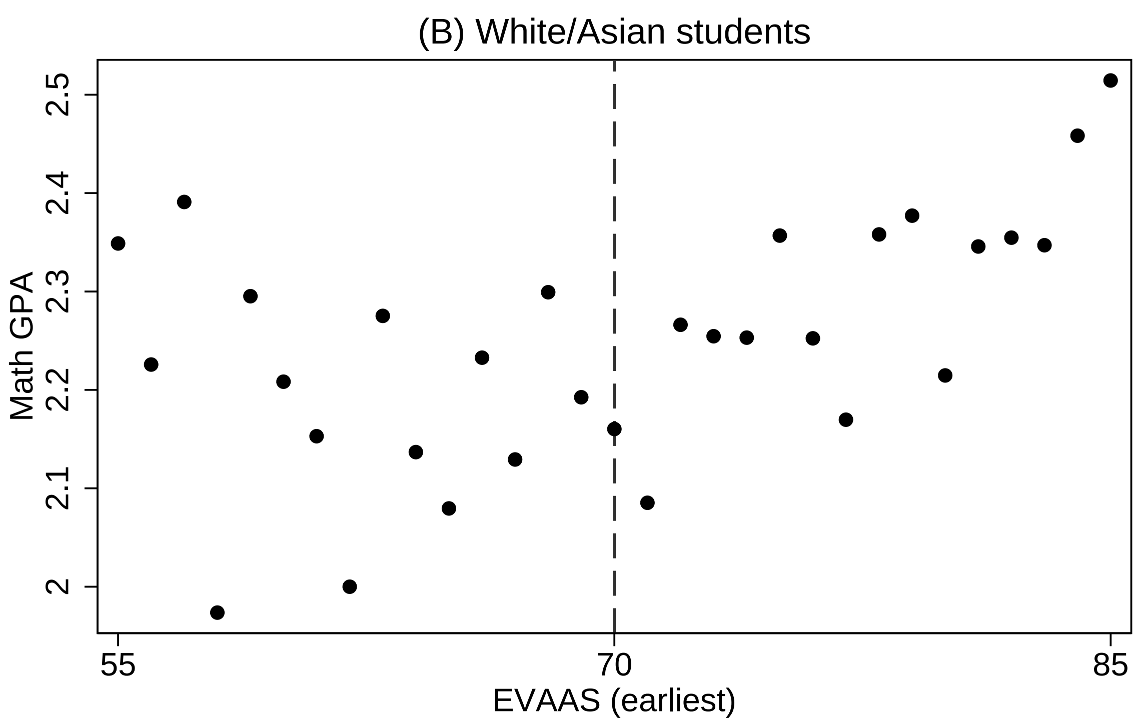


Figure A.9: Math Z-Score, by Income

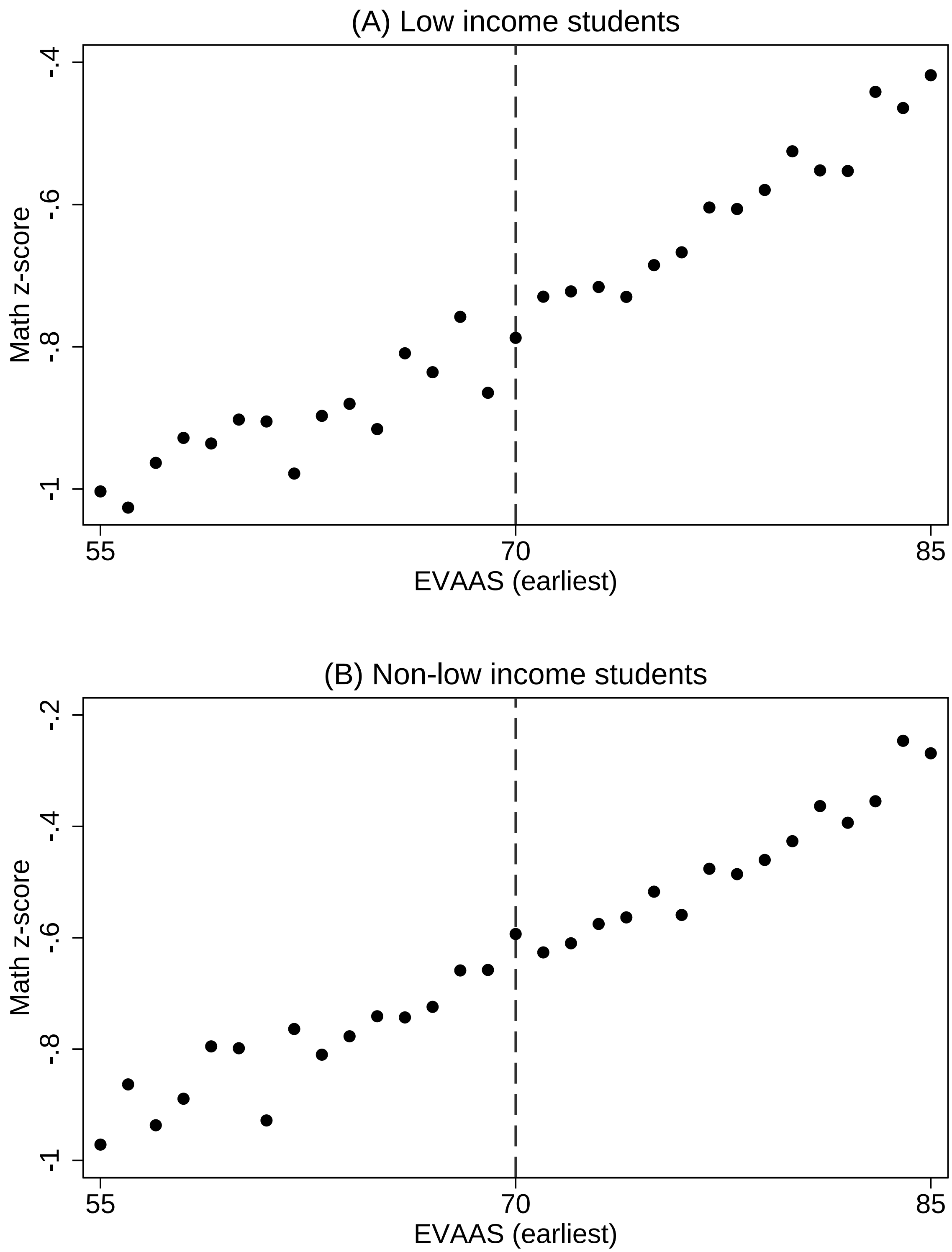


Figure A.10: Math Z-Score, by Race

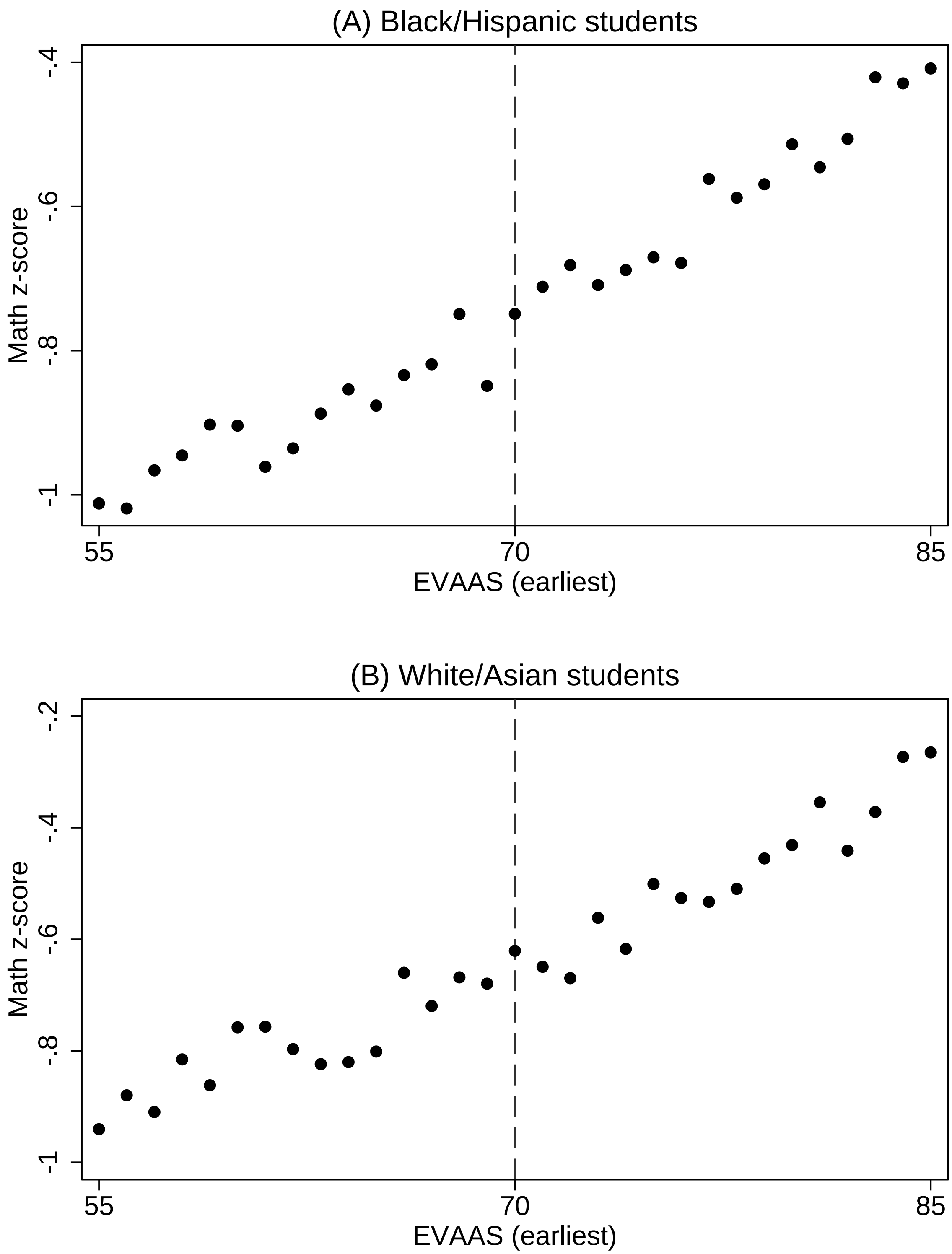


Table A.1: Robustness Checks, 2010-2013 Cohorts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First		Math GPA			Math z-score	
	stage	All	Female	Male	All	Female	Male
Bandwidth = 5	0.137***	-0.683*	-1.118*	-0.572	0.464*	0.250	0.683
	(0.027)	(0.401)	(0.614)	(0.820)	(0.254)	(0.310)	(0.433)
Bandwidth = 10	0.137***	-0.578**	-1.445**	-0.014	0.204	-0.053	0.313
	(0.021)	(0.293)	(0.562)	(0.358)	(0.171)	(0.264)	(0.219)
Bandwidth = 15	0.134***	-0.541**	-1.400***	0.030	0.142	-0.186	0.343*
	(0.019)	(0.274)	(0.386)	(0.324)	(0.119)	(0.186)	(0.179)
Bandwidth = 20	0.133***	-0.704***	-1.349***	-0.186	0.001	-0.232	0.176
	(0.016)	(0.273)	(0.298)	(0.348)	(0.115)	(0.157)	(0.181)
Bandwidth = 25	0.132***	-0.896***	-1.439***	-0.430	-0.129	-0.307**	0.019
	(0.016)	(0.252)	(0.278)	(0.335)	(0.124)	(0.152)	(0.185)
BW = 15, controls	0.130***	-0.609**	-1.645***	0.027	0.103	-0.305	0.347**
	(0.018)	(0.275)	(0.475)	(0.315)	(0.123)	(0.212)	(0.177)
IK Bandwidth	0.134***	-0.513*	-1.429***	0.068	0.166	-0.150	0.322*
	(0.018)	(0.274)	(0.439)	(0.326)	(0.138)	(0.211)	(0.188)
BW	16.7	13.3	13.0	12.2	12.3	12.7	12.2

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$). Column 1 shows first stage estimates of the relationship between eligibility for math acceleration and the fraction of middle school years spent in accelerated coursework. The remaining columns show instrumental variables estimates of the impact of math acceleration on various outcomes. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. The first five rows use bandwidths from 5 to 25. The sixth row uses a bandwidth of 15 and includes controls for gender, race, age, poverty, special education and LEP status. The seventh row uses the smaller of the two Imbens-Kalyanaraman optimal bandwidths generated by the first stage and reduced form specifications. The sample includes the 2010-2013 cohorts.

Table A.2: First Stage Impacts, By Gender, Poverty and Race

	(1) 2010-2013 cohorts	(2) 2010 cohort	(3) 2011 cohort	(4) 2012 cohort	(5) 2013 cohort
(A) Gender					
Male * Eligible	0.160***	0.087***	0.131***	0.203***	0.314***
	(0.018)	(0.031)	(0.037)	(0.052)	(0.068)
Female * Eligible	0.111***	0.063*	0.051	0.171***	0.259**
	(0.028)	(0.032)	(0.036)	(0.049)	(0.098)
Female	0.035	0.004	0.037	0.017	0.128*
	(0.023)	(0.029)	(0.027)	(0.047)	(0.074)
p	0.08	0.56	0.08	0.67	0.58
N	16,010	4,910	4,956	4,010	2,134
(B) Income					
Non-poor * Eligible	0.140***	0.124**	0.049	0.185***	0.265***
	(0.023)	(0.049)	(0.039)	(0.046)	(0.087)
Poor * Eligible	0.129***	0.042*	0.113***	0.187***	0.285***
	(0.026)	(0.024)	(0.035)	(0.049)	(0.090)
Poor	-0.030	-0.001	-0.044	-0.040	-0.089
	(0.025)	(0.027)	(0.031)	(0.058)	(0.080)
p	0.75	0.14	0.18	0.97	0.86
N	16,010	4,910	4,956	4,010	2,134
(C) Race					
White/Asian * Eligible	0.097***	0.092**	0.011	0.099**	0.264***
	(0.025)	(0.035)	(0.046)	(0.045)	(0.078)
Black/Hispanic * Eligible	0.152***	0.066**	0.124***	0.236***	0.286***
	(0.024)	(0.031)	(0.033)	(0.047)	(0.094)
Black/Hispanic	-0.027	0.012	-0.045	-0.040	-0.085
	(0.026)	(0.022)	(0.046)	(0.062)	(0.070)
p	0.10	0.57	0.04	0.05	0.85
N	16,010	4,910	4,956	4,010	2,134

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). First stage estimates show the impact of eligibility for acceleration on the fraction of middle school years spent in accelerated math coursework. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. These replicate the regressions from the top row of Table 3, interacting the independent variables with indicators for gender or race. Also shown is a p-value from an F-test of the equality of the two interaction coefficients shown. The sample includes the 2010-13 cohorts.

Table A.3: Heterogeneity of Impacts, By Income and Race

	(1) Math GPA	(2) At least D	(3) At least B	(4) Math z-score
(A) Income				
Non-poor * Accelerated	-1.201**	-0.196***	-0.488**	0.227
	(0.490)	(0.071)	(0.229)	(0.239)
Poor * Accelerated	-0.285	-0.073	-0.234	0.058
	(0.379)	(0.090)	(0.172)	(0.162)
Poor	-0.704**	-0.094*	-0.248*	-0.009
	(0.307)	(0.047)	(0.131)	(0.132)
μ (Non-poor)	2.35	0.97	0.46	-0.68
μ (Poor)	2.02	0.92	0.35	-0.82
p	0.18	0.26	0.39	0.55
N	16,010	16,010	16,010	15,850
(B) Race				
White/Asian * Accelerated	-0.450	-0.083	-0.083	0.007
	(0.613)	(0.141)	(0.325)	(0.331)
Black/Hispanic * Accelerated	-0.554**	-0.114	-0.361**	0.192
	(0.264)	(0.074)	(0.136)	(0.143)
Black/Hispanic	-0.074	0.001	0.046	-0.123
	(0.278)	(0.073)	(0.146)	(0.166)
μ (White/Asian)	2.21	0.95	0.42	-0.69
μ (Black/Hispanic)	2.11	0.94	0.38	-0.80
p	0.87	0.86	0.40	0.63
N	16,010	16,010	16,010	15,850

Notes: Heteroskedasticity robust standard errors clustered by initial middle school are in parentheses (* p<.10 ** p<.05 *** p<.01). Instrumental variables estimates show the impact of the fraction of middle school years spent in accelerated math courses on various measures of course grades and test scores, where that fraction is instrumented by eligibility. The coefficients shown are generated by local linear regression using a triangular kernel of bandwidth 15, including cohort-by-school-by-grade fixed effects. That specification is then interacted with indicators for each demographic subgroup. Below each coefficient are the subgroup means of the outcome variable among students just below the threshold. Also shown is a p-value from an F-test of the equality of the two interaction coefficients. The sample includes the 2010-13 cohorts.