NBER WORKING PAPER SERIES

CAN GIFTED EDUCATION HELP HIGHER-ABILITY BOYS FROM DISADVANTAGED BACKGROUNDS?

David Card
Eric Chyn
Laura Giuliano

Working Paper 33282
http://www.nber.org/papers/w33282

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2024

We are extremely grateful to Cynthia Park for assistance in accessing and interpreting the data used in this study as well as information about the District's policies and practices, and to Tatiana Reyes for extraordinary research assistance. We also thank seminar participants at Brigham Young University, Duke University, McGill University, Queen's University, RWI Essen, Simon Frazer University, UC Merced, UC Riverside, UC Santa Barbara, University of Nevada-Las Vegas, University of Utah, University of Virginia, and the Upjohn Institute. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through grant R305C200012 to the National Center for Research on Gifted Education. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Can Gifted Education Help Higher-Ability Boys from Disadvantaged Backgrounds?
David Card, Eric Chyn, and Laura Giuliano
NBER Working Paper No. 33282
December 2024
JEL No. I21, I24

ABSTRACT

Boys are less likely than girls to enter college, a gap that is often attributed to a lack of non-cognitive skills such as motivation and self-discipline. We study how being classified as gifted – determined by having an IQ score of 116 or higher – affects college entry rates of disadvantaged children in a large urban school district. For boys with IQ's around the cutoff, gifted identification raises the college entry rate by 25-30 percentage points – enough to catch up with girls in the same IQ range. In contrast, we find small effects for girls. Looking at course-taking and grade outcomes in middle and high school, we find large effects of gifted status for boys that close most of the gaps with girls, but no detectable effects on standardized tests scores of either gender. Overall, we interpret the evidence as demonstrating that gifted services raise the non-cognitive skills of boys conditional on their cognitive skills, leading to gains in educational attainment.

David Card
Department of Economics
549 Evans Hall, #3880
University of California, Berkeley
Berkeley, CA 94720-3880
and NBER
card@econ.berkeley.edu

Eric Chyn
The University of Texas at Austin
Department of Economics
2225 Speedway, Stop C3100
Austin, TX 78712-1690
and NBER
eric.chyn@austin.utexas.edu

Laura Giuliano
Department of Economics
University of California, Santa Cruz
1156 High Street
Santa Cruz, CA 95064
and NBER
lgiulian@ucsc.edu

Boys are less likely to enter college than girls (Jacob, 2002; Goldin et al., 2006). The gap is concentrated among children from disadvantaged backgrounds: recent data from the Postsecondary Student Aid Study show that the ratio of males to females among college freshmen is close to 1 for children from the top income quartile, but falls to only about 60% for children from the bottom quartile (see Appendix Figure A1).[1] Even more striking is the disparity for disadvantaged boys with higher cognitive ability. Figure 1 shows college entry rates by gender and participation in the free- and reduced-price lunch (FRL) program for children with differing levels of cognitive ability, as measured by IQ and standardized tests in second grade.[2] At low levels of ability, the differences between the four groups are small. But at higher levels, low-income boys lag far behind the other groups.

What is going wrong for these boys? Jacob (2002) and Becker et al. (2010) argue that the gender gap in college entry arises from boys' lack of *non-cognitive* skills like self-discipline and grit, and show that controlling for measures of these skills can explain most of the difference between girls and boys. Likewise, Heckman and Rubinstein (2001) and Heckman et al. (2006) argue that socioeconomic gaps in schooling attainment are driven in large part by the under-development of non-cognitive skills. Indeed, the model estimated by Heckman et al. (2006, Figure 21) shows that higher cognitive skills have little effect on college completion rates unless they are accompanied by higher non-cognitive skills. These two strands of work suggest that disadvantaged boys face a *double penalty* in non-cognitive skill formation that could be especially important for those with higher cognitive skills.

There is less evidence on whether K-12 education policies can improve non-cognitive skills and help close the gender gap in college entry. Jackson (2018) shows that certain 9th-grade teachers can increase non-cognitive skills and boost high school completion rates, though he does not compare girls and boys. Rose et al. (2022) show that exposure to teachers who raise student attendance and other measures of non-cognitive skill lowers subsequent criminality, with

---

[1] Many recent papers have emphasized the problems faced by boys from disadvantaged backgrounds. See, for example, Bertrand and Pan (2013) Chetty et al. (2016) and Autor et al. (2019).

[2] These data are from the school district we study and include children who were in 2nd grade in 2008-2010. We measure disadvantage by participation in the free or reduced price lunch (FRL) program. We measure cognitive ability by a combination of math and reading scores on the 2nd-grade SAT test and the Naglieri Non-verbal Ability Test (an IQ-like test used in the district to screen for gifted children) and scale ability by the predicted college entry rate of non-disadvantaged girls. There are only small mean differences by gender in the distributions of ability, though boys have somewhat fatter tails than girls. Goldin et al. (2006) report that distributions of IQ among high school juniors in the Wisconsin Longitudinal Survey are almost identical for boys and girls.

similar effects for girls and boys. In a randomized intervention aimed at higher-ability children in Italy, Carlana et al. (2022) find that a middle school tutoring and counseling program raises the fraction of high-achieving immigrant boys who choose an academic high school track, closing the gap with their female counterparts.[3]

A long-established constellation of gifted programs in the U.S. – serving about 3 million students per year – is designed in part to ensure that children with high cognitive abilities remain engaged in school and progress to college (VanTassel-Baska, 2018).[4] In this paper, we study the effects of being identified as gifted in elementary school on subsequent test scores, courses and grades, and on the probability of entering college, for disadvantaged children in a large urban school district (hereafter, "the District").[5] To qualify for gifted status in the District, English language learners (ELL's) and FRL participants must have an IQ score of 116 points or higher (roughly the 84th percentile in the national distribution). This comparatively generous eligibility standard, coupled with a universal testing program to identify high-ability students, means that the District serves a relatively large number of disadvantaged students in its gifted program. The strict IQ threshold for gifted status also provides the basis for a regression discontinuity (RD) evaluation of the impacts of being identified as gifted.[6]

Focusing on ELL/FRL students who had a first IQ test by the end of 5th grade, scored 106-124 points, and remained in the District for 7 more years – a sample of about 3,500 students – we show there is a large discontinuity in the probability of being classified as gifted at the 116-point threshold.[7] Once identified, gifted students retain their status through the end of high school and receive a package of services in subsequent years. These include a state-mandated Education Plan (EP) specifying instructional goals, and biannual meetings to review students'

---

[3] Broadly consistent with the results, though operating at a much later stage, Canaan, Fischer, Mouganie & Schnorr (2024) find that a low-touch counseling intervention for college students on probation leads to improvements in GPA, completion, and earnings – with benefits concentrated among low-SES students, males, and STEM majors.

[4] Only a handful of studies examine the causal effects of these programs. Bui et al. (2014) study the impacts of gifted status on test scores – like us, they find no significant effects (see below). We discuss related studies below.

[5] Our data sharing agreement with the District requires that it is not specifically identified. The District is located in Florida.

[6] Non-disadvantaged students face an IQ eligibility threshold of 130 points. Results available on request suggest that gifted identification has no significant impact on college entry for these students, though as noted by Card and Giuliano (2014), there is substantial manipulation of IQ scores around the 130-point threshold.

[7] Nearly all students who are ever classified as gifted in the District have an IQ test by the end of 5th grade, and ~80% are IQ-tested before 4th grade. We include those tested in grades 4 and 5 to increase our sample size. We focus on students who remain in the District until the nominal high school completion time so we can measure their middle and high school outcomes and college attendance. As discussed below, we find no evidence of a discontinuity in the probability of remaining in the District.

progress and identify options at key transition points.[8] In addition, gifted students in the District are assigned to separate gifted/high achiever (GHA) classrooms for 4th and 5th grades and can receive higher priority for advanced-track courses, particularly in middle school.

We evaluate the impacts of gifted status by 5th grade on college entry 7 years later (i.e., "on time") and on a wide range of intermediate outcomes in elementary, middle and high-school. A concern for our analysis is that psychologists have some discretion in scoring IQ tests, leading to bunching at the threshold for gifted eligibility. We address this with a standard "donut RD" approach (Barreca et al., 2011).[9] To assess the robustness of our results, we present RD estimates using a range of bandwidths for all outcomes, and we summarize the impacts on college entry using a two-dimensional grid of bandwidths to the right and left of the IQ cutoff.

Our main finding is that for disadvantaged boys with IQ's near the 116-point cutoff, being identified as gifted leads to a 25-30 percentage point (ppt) increase in the probability of completing high school on time and entering college the following year, from just under 50% for boys who narrowly miss gifted status to around 75% for those who meet the threshold – comparable to the rate for similar-IQ girls with scores just under the threshold. In contrast, we find relatively small, statistically insignificant effects (~5 ppt) on the college entry rate of girls.

We go on to explore the effects of gifted status on a wide range of intermediate outcomes, using detailed information on course selections, course grades, and standardized test scores. We find no effects on statewide standardized test scores or on PSAT scores for either boys *or* girls, suggesting that gifted services have little or no impact on cognitive skills. Looking at outcomes that are widely interpreted as markers of non-cognitive skill, however, we find large positive effects for boys. Specifically, we find a 25-ppt increase in the probability that boys enter an accelerated math program in 6th grade, a nearly 30-ppt increase in the probability of completing Algebra 1 before 9th grade, and a doubling of the number of AP classes taken in high school. We also measure a 0.4-point increase in average grades in high school math (on a standard 4-point scale), and relatively large impacts on the quality of classroom peers – indicating that boys who meet the gifted threshold are more likely to take advanced-level classes.

---

[8] EP's – which are analogous to the plans developed for Special Needs children – are a central feature of gifted programming in most states (Zirkel, 2005).

[9] We also use only the first reported IQ score to avoid concerns about selective re-testing.

In all these domains, passing the gifted threshold allows boys to catch up (or nearly catch up) to girls with similar IQ's, whereas there is little or no effect for girls. We offer a simple explanation for this pattern, based on the hypotheses that cognitive and non-cognitive skills are complements in the education production function (e.g., Heckman et al. 2006), and that gifted services boost the non-cognitive skills of students, with little or no effect on their cognitive skills. The non-cognitive boost is important for disadvantaged boys with higher IQ's, whose non-cognitive skills are low relative to their cognitive skills, particularly in the math domain.[10] Our results suggest that gifted services raise their non-cognitive skills to levels comparable to those of girls with similar IQ scores, leading to increased engagement in middle and high school and higher college entry rates. In contrast, we conjecture that even in the absence of gifted services, most girls have non-cognitive skills that are aligned with their cognitive skills. Thus, a boost in non-cognitive skills with no change in cognitive skills will have relatively little effect on their grades, participation in advanced courses, or college entry.

We provide some direct evidence for this explanation using a combination of available measures of non-cognitive skill for students in third grade (prior to entry to specialized classrooms for gifted and high-achieving students in 4th grade).[11] For children of both genders with higher non-cognitive skills, we find small effects of gifted status on college entry. For boys and girls with lower non-cognitive skills, however, we find relatively large positive treatment effects of gifted status – suggesting that these services are especially valuable for students who were less engaged with schooling in early elementary grades regardless of gender. But boys have substantially lower non-cognitive scores, so the average effect is higher for them.

Our paper contributes to four distinct literatures. The first is a body of work addressing the gender gap in college entry/completion (e.g., Jacob, 2002; Charles and Luoh, 2003; Goldin et al., 2006; Vincent-Lancrin, 2008; Becker et al., 2010; Fortin et al., 2015). A second and related literature examines the interactions between family background and gender, noting that disadvantaged backgrounds are associated with particularly bad outcomes for boys (e.g., Bertrand and Pan, 2013; Chetty et al., 2016; Autor et al., 2019; Aucejo and James, 2019). Our work builds on both literatures, focusing on the gender gap in college entry for ELL and FRL

---

[10] We use standardized test scores to measure domain-specific cognitive skills.

[11] As discussed below, most gifted students are identified in 3rd grade and receive some limited services in that grade, but we believe the impact of these services is small.

children with relatively high cognitive ability. Like only a few previous studies – most notably, Carlana et al. (2022) – we contribute by studying the causal effects of a package of services that narrows the gender gap for higher-ability but disadvantaged children.

A third, much smaller, literature analyzes the effects of gifted programming on student outcomes. Bui et al. (2014) and Card and Giuliano (2014) use RD designs based on gifted eligibility rules to examine impacts on standardized test scores 1-2 years after program entry. Like us, they find no significant impacts on test scores. Bui et al. (2014) also use a lottery design to study effects of being admitted to magnet schools, again finding no effects on math or reading scores, but some gains in science.[12] Booij et al. (2016) and Lavy and Goldstein (2022) analyze high-school gifted programs in the Netherlands and Israel and find no effects on test scores or college entry (though college rates are already high in their samples).[13] We contribute to this literature by providing the first design-based study of the longer-term effects of gifted identification in a U.S. setting. Our focus on disadvantaged students around the 85th percentile of IQ's departs from most of the gifted education literature (see e.g., Subotnik et al., 2011), which focuses on children in the extreme upper tail of ability (i.e., the top 5% or even 1%).

Finally, our paper is related to a few other studies that look at advanced-track programs for high-achieving students in elementary grades.[14] Cohodes (2020) examines an accelerated curriculum program in 4th-6th grades for students who score highly on a 3rd-grade standardized exam and finds no effects on standardized test scores but a positive effect on college entry for Black and Hispanic students. She does not report results by gender. In previous work focused on advanced academic programs in the District, Card and Giuliano (2016b) examined the effects of assignment to GHA classrooms in grades 4-5 for non-gifted students. These students are selected on the basis of their previous year's test scores to fill open seats in GHA classrooms at schools with too few gifted students. In contrast to the null effects on test scores that we find for gifted students, Card and Giuliano (2016b) find positive effects for non-gifted, high-achieving students that are concentrated among Blacks and Hispanics and significant for both boys and girls.

---

[12] An earlier study by Bhatt (2011) uses an instrumental variables approach and finds some evidence of a positive effect of being identified as gifted on test scores. As noted by Bui et al. (2014) her design is fragile. An unpublished study by Murphy (2009) also finds no systematic effects of gifted programs on test scores.

[13] Booij et al. (2016) find positive effects on course grades; Lavy and Goldstein (2022) find effects on double majoring and pursuing advanced degrees.

[14] More tangentially related are studies of selective high schools and high school tracks. Abdulkadiroglu et al. (2014) and Dobbie and Fryer (2014) find no effects of selective high schools in Boston and New York City on relatively short-run outcomes.

A key distinction between the gifted program we study and the test-based "tracking" programs studied elsewhere is the use of IQ scores to identify students as gifted. Indeed, many gifted students in our sample – especially those who just meet the 116 IQ threshold used for disadvantaged students – are far from the top of their class in grades or test scores.[15] Also unlike the non-gifted high achievers studied by Card and Giuliano (2016b), gifted students in the District receive an EP and support services throughout their schooling careers.

The next section presents a simple model of educational attainment that emphasizes the complementarity of cognitive and non-cognitive skills and provides an interpretative framework for our empirical analysis. We then describe the District's gifted program and our data set, which combines student records from the District with college enrollment data from the National Student Clearinghouse. Next, we document the first-stage relationship between IQ scores and gifted status. We then turn to RD-based estimates of the effect of gifted status on our main outcome, on-time graduation and college entry, including an investigation into the differential effects by gender. We follow with an analysis of intermediate outcomes, focusing on patterns for boys versus girls and distinguishing between proxies for cognitive vs. non-cognitive skills. The final section of the paper offers an interpretation of our findings and their implications for policy.

I. A Simple Framework

Heckman et al. (2006) propose an education production function in which schooling attainment (y) is related to two complementary factors: student cognitive ability (C) and student non-cognitive ability (N).[16] To simplify their model while emphasizing the complementarity between cognitive and non-cognitive skills, assume that:

$$y = \min[\, C \,,\ N \,] + e \tag{1}$$

where e incorporates all other factors. In this setup, an education intervention can affect attainment by raising C or N (or both). The impact of an intervention will also depend crucially on the distribution of C and N in the target population. Here, we study the effects of gifted

---

[15] In Appendix Table A1, we present a side-by-side comparison of the compliers in our RD design for disadvantaged gifted students and the compliers in the RD design used in Card and Giuliano (2016b) for non-gifted high achievers. The two groups differ on several dimensions, including race/ethnicity, FRL status, neighborhood income, classroom composition, and baseline measures of cognitive and non-cognitive ability.

[16] There is a long history of education production function modeling in economics, starting with Bowles (1970). One strand of this literature focuses on teacher value-added – see e.g., Hanushek (2020). Another focuses on the production of cognitive skill, e.g., Todd and Wolpin (2003). We follow the approach of Heckman et al (2006) in modeling the production of educational attainment (e.g., grade completion or college entry) and focusing on the role of different types of skills.

services on students with IQ scores (measured in early elementary grades) close to the gifted eligibility threshold. As we show below, these services have little or no effect on statewide standardized test scores and PSAT scores – outcomes that are usually interpreted as measures of cognitive skill (e.g., Jackson 2018). Under the assumptions in equation (1) this means that the impacts of gifted services will be driven entirely by gains in non-cognitive skills among students whose non-cognitive skills are low relative to their cognitive skills.

Building on Jacob (2003) and Becker et al (2010), assume that girls with cognitive skills at the level of the IQ cutoff for gifted eligibility (i.e., $C = C_0$) have non-cognitive skills that are closely aligned with their cognitive skills:

$$E[N \mid C = C_0, girl] = C_0,$$

whereas in the absence of gifted services, boys with IQ's in the same range have a relative deficit of non-cognitive skills:

$$E[N \mid C = C_0, boy] = (1 - d)C_0,$$

where $d > 0$. In this case, even if gifted services lead to a gender-neutral boost ΔN in non-cognitive skills, the effect on girls' educational attainment will be relatively small, while the effect on boys will be larger, and potentially nearly as large as ΔN.[17]

A useful extension of model (1) is to allow for different domains of outcomes and skills. For example, we could let y_M and y_L be two domain-specific intermediate outcomes (e.g., grades in math and grades language arts) that are substitutes in the production of college enrollment. In turn, suppose each outcome y_D is related to domain-specific cognitive and non-cognitive skills C_D and N_D, $D \in \{M, L\}$, following (1). It is plausible that some students have relatively high cognitive skills in one domain (e.g., math), but lack the relevant non-cognitive skills to achieve at their "full potential" in that subject area (e.g., $N_C = N_M \leq C_L < C_M$). In this scenario, an intervention that boosts N_C and N_M equally would tend to have a larger effect on y_M; moreover, the same effect on y_M could be achieved by targeting only N_M. As we discuss below, disadvantaged boys who barely pass the IQ threshold for gifted status (i.e., the complying boys) have higher standardized test scores in math than in reading, and gifted services have relatively

---

[17] If $\Delta N > 0$, the gain in expected attainment for a student with skills (C_0, N) is $G = \min\,[\Delta N, \max\,(C_0 - N, 0)]$. Assuming the distribution of N conditional on C is symmetric and relatively tight, half of girls have $N > C_0$ and get no boost in expected attainment, while those with $N < C_0$ only benefit by the gap between their non-cognitive and cognitive skills, which will be small on average. Under these same conditions, most boys will benefit at least a little from the gain in non-cognitive skills, and those with $N < C - \Delta N$ will benefit by $G = \Delta N$.

large impacts on their math-related outcomes in middle and high school. These findings suggest that for many boys, gifted services provide the non-cognitive boost they need to achieve outcomes consistent with their cognitive abilities in math.

Our main outcome of interest is on-time college entry, which is arguably determined by combination of cognitive and non-cognitive skills in one or more subject domains. We also observe a wide set of intermediate outcomes, including subject-specific GPA's in high school and the number of AP courses completed by 12th grade. In prior studies of non-cognitive skill gaps (e.g. Jacob 2002; Cornwell, Mustard and Van Parys, 2013), grades and course selections are interpreted as markers of non-cognitive skill *when cognitive skills are held constant*.[18] Since our models condition on IQ, we follow this interpretation and refer to these outcomes as measures of non-cognitive skill, though we believe that most of these outcomes depend on a combination of non-cognitive *and* cognitive skills. As we show below, we find generally similar treatment effects for course taking and GPA outcomes as for college entry – a pattern that is consistent with the hypotheses that there is strong complementarity between cognitive and non-cognitive skills and a non-cognitive skill deficit among boys that is partly remediated by gifted services.

II. Background and Data

The District is one of the 10 largest school districts in the country, serving around 20,000 students per grade. Its student body is racially diverse (see Table 1); about one-half are eligible for the FRL program. During our sample period, the District operated 140 larger elementary schools offering Kindergarten to 5th grade, 80 conventional middle schools offering grades 6-8, and 70 mainstream high schools. It also offered a wide variety of charter and special service schools. The District's policies for gifted students, including eligibility standards and services, are regulated by states laws and have been relatively stable since 2002, with some exceptions noted below.

a. Gifted Identification

Most gifted children in the District are identified between 2nd and 5th grades, with the modal student entering in 3rd grade (see Appendix Figure A2). Once identified, gifted students

---

[18] Cornwell et al. (2013) directly link early-grade gender gaps in non-cognitive measures like classroom engagement to differences in teacher-assigned grades between boys and girls with similar test scores.

retain their status throughout their schooling careers through the end of high school. To be eligible for gifted status, state law requires students to achieve a minimum score on a standard IQ test (administered by a licensed psychologist). The law specifies a threshold of 130 points for non-disadvantaged students, but allows districts to select a lower "Plan B" threshold for FRL and ELL students, which the District sets at 116 points. Since IQ testing is costly ($1,000 or more for a test by a private psychologist), nearly all lower-income children are tested by District psychologists.

Once a student is found to meet the IQ threshold, the school's Exceptional Student Education (ESE) specialist meets with the student, parents and teachers to determine whether additional gifted criteria are met, including the "need for a special instructional program." The final determination weighs a number of factors, including test scores and behavioral indicators. Notably, the District's policy is intentionally flexible with respect to traditional criteria like leadership, creativity and motivation, so as not to eliminate gifted "underachievers" from the gifted program.

The District also promotes diversity in its gifted population through a "universal screening" program. Traditionally, students are referred for gifted evaluation through a process that depends on parents and teachers, leading to concerns about the under-representation of minorities and lower-income children (Ford, 1998). In 2005, the District began administering an in-class, non-verbal ability test to all 2nd graders; it then used cutoffs on this test (matching the IQ thresholds for gifted status) to generate automatic referrals for IQ testing. As shown in Card and Giuliano (2016a), this led to an immediate rise in the fraction of "Plan B-eligible" children who received an IQ test (from 13% to 24%) and who were identified as gifted by the end of 3rd grade (from 1.4% to 4.3%). Among all children eligible for Plan B in the cohorts covered in our schooling data, 26% had an IQ test and 3.6% were identified as gifted by the end of 5th grade.[19]

b. Gifted Services

Once identified, gifted students are eligible to receive gifted services through the end of high school. The starting point for these services is an EP, which is developed by an ESE specialist along with input from teachers and parents. State guidance instructs that the EP should

---

[19] Of the 10 cohorts in our sample, 7 were exposed to universal screening; however, the rates of IQ testing varied across cohorts due to a scaling back of the program during the Great Recession (2008-2011), followed by its re-institution in 2012 using a different screening test.

identify: (i) the student's needs related to his or her area of giftedness (e.g., exceptional math ability) and (ii) the services that will be provided to ensure the student will progress appropriately. While gifted EP's are required to specify curriculum goals, they can also include social/emotional and "independent functioning" goals (e.g., related to homework completion). Indeed, state guidelines encourage staff to consider what is "the most important skill or behavior that the student needs in order to make progress commensurate with his/her abilities" (Florida Department of Education 2006).

Prior to 4th grade, gifted students typically receive a few hours per week of individualized instruction based on the learning goals in their EP. Starting in 2004 – and affecting 8 of the 10 cohorts in our analysis sample – the District mandated separate classrooms in 4th and 5th grades for all gifted students, with any extra seats filled by each school's highest-scoring students on the prior year's state-wide tests ("high achievers"). We refer to these as gifted/high achiever (GHA) classrooms. GHA classrooms follow the same curriculum and use the same textbooks as in regular classes, although gifted students (unlike the non-gifted high achievers in the same classrooms) continue to pursue individualized goals as assigned in their EP's.

Most of the District's middle schools also offer separate gifted/high achiever classes for language arts in grades 6-8. In addition, all middle schools offer a 3-year accelerated math program, known as "Great Explorations in Math" (GEM), which combines pre-algebra in 6th grade, algebra in 7th grade, and geometry in 8th grade. Students with a minimum score on the 5th-grade statewide math test are eligible to enter GEM; continuation in later years is based on course grades. Gifted students may receive priority for filling extra seats in a GEM class if there are not enough students scoring above the 5th-grade cutoff. In high school, gifted students often take honors-level classes (such as calculus) or participate in specialized programs such as the International Baccalaureate. Many also take advanced placement (AP) courses.

Finally, gifted students have access to specialized teachers and counselors throughout their schooling careers. GHA classrooms at all grade levels are assigned teachers who have completed training in gifted education, including content related to disadvantaged populations and strategies to meet children' affective, social and emotional needs. Gifted students meet with their ESE specialist at least once every two years to review and update the goals in their EPs. In addition, at the transitions to middle and high school, students and parents meet with the ESE specialists of the sending and receiving schools. One goal of these meetings is to help students choose appropriate courses to match their abilities. The findings of Carlana et al. (2022) suggest that this guidance may be especially helpful for disadvantaged students who lack college-educated role models.

c. *Data Sources*

We use student-level data from the District's centralized record system for children who were enrolled in 5th grade in the 2003-2012 academic years. These records contain information on each student's gender, race/ethnicity, FRL/ELL status, and gifted status. They also include state-wide achievement test scores in reading and math for grades 3-8; disciplinary logs from grades 3-8; Stanford Achievement Test scores from 2nd grade (for 8 of our sample cohorts); scores from the gifted screening test that was adopted in 2005 and administered in 2nd grade (for 5 of our cohorts); and PSAT scores for all those who took the exam in high school (around 92% of our sample). Courses enrollments are available for grades 3-12, and course grades are available for middle and high school.

Information on college enrollment is reported to the District by the National Student Clearinghouse (NSC) and is available for all students who completed high school in the District. For a subset of grades and cohorts, we can also link student responses from various District-administered surveys. Finally, we link these records to IQ test scores for all students in our sample cohorts who were tested by the District or submitted a score on their own.

d. Analysis Sample and Descriptive Statistics

We derive our main analysis sample in three steps. First, we identify students who had a first IQ test administered by 5th grade, were ELL or FRL at the time of the test, scored 100-128 points, were enrolled in the District by 4th grade or earlier, and did not repeat 5th grade. Second, we require that the student stayed in the District for at least seven years after 5th grade, had non-missing test scores in grades 5-8, and had non-missing data on courses and grades in middle and high school.[20] This second step, which reduces the sample by around 25%, ensures that we can observe information on high school graduation and college entry. Finally, for our main analysis sample, we further limit attention to students with IQ's in the range of 106 to 124 points.

---

[20] Due to missing data in one year of our sample, high school course records are incomplete for 4 of our 10 cohorts. We include these cohorts in our analysis of high school GPA's by averaging over courses in available years; however, we omit these cohorts for our analysis of AP course-taking in high school. Our conclusions regarding high school GPA are robust to models fit to the restricted set of cohorts, and all of our results are robust to specifications that include cohort fixed effects.

A potential concern with our sample is that students who achieve gifted status may be more (or less) likely to remain in the District through the end of high school, leading to selection biases. To address this concern, we use our "step 1" sample (and gender-specific sub-samples) to estimate RD models for the effect of passing the 116 IQ threshold on the probability of being in our "step 2" sample. The estimates are presented in Appendix Table A2. Reassuringly, these estimates are uniformly small and statistically insignificant.

Table 1 presents summary statistics for students in our main analysis sample, focusing on predetermined characteristics of students and their schools.[21] As a benchmark, column 1 reports statistics for all students who were in 5th grade in the District in 2003-2012. The district is racially and ethnically diverse, with roughly equal shares of White, Black and Hispanic students. About one-half were FRL-eligible and 8% were English language learners (ELL) – rates that are comparable to other large urban school districts.

We also summarize cognitive and non-cognitive skills in early grades. We measure cognitive skills in reading and math using (standardized) test scores on 2nd-grade Stanford Achievement tests (SAT's) and on 3rd-grade statewide tests. Since our data do not include information from elementary school report cards, we measure non-cognitive skills using two behavioral variables that are available for most 3rd-graders: (i) the number of disciplinary actions and (ii) student responses to a survey asking whether they "enjoy learning" at their school. Since each has limited variation, we combine these two variables and classify students as having high non-cognitive skills if they had no disciplinary actions and "strongly agreed" that they enjoyed learning at school. As show in column 1, 60% of the District-wide sample have high non-cognitive skills by this measure.

Column 2 describes our main analysis sample of Plan B students with a first-reported IQ test in the 106-124-point range who were still in the District seven years after 5th grade. Consistent with their above-average IQ's, these students are relatively high-achieving, with mean test scores in 2nd and 3rd grade that are 0.5-0.6 standard deviations (σ's) above the District-wide average. They are also more likely to be minorities, have a high rate of FRL participation, and attended elementary schools with higher fractions of FRL participants and slightly lower standardized test scores than the sample in column 1. Interestingly, this high-IQ sample is only slightly more likely than the District-wide sample to have high non-cognitive skills. Columns 3

---

[21] Outcome variables are summarized in Appendix Table A3 and discussed later.

and 4 show characteristics of the analysis sample separately for boys and girls. There are relatively small differences between them, though girls have somewhat higher (by about 0.25 σ's) test scores in reading and are 8 ppts (14%) more likely have a high non-cognitive score.[22] Columns 5 and 6 of the table show the characteristics of compliers for our RD analysis of gifted identification, which we discuss below.

III. Empirical Strategy

We adopt a fuzzy RD approach for analyzing the effect of gifted status on student outcomes, estimating a locally linear first-stage model for gifted status that has a jump at the threshold for Plan B gifted eligibility and parallel reduced-form models for outcomes like college enrollment. Under standard assumptions (e.g., Hahn, Todd and van der Klaauw, 2001), the ratio of the jump in the outcome at 116 points to the jump in the fraction of gifted students provides a causal estimate of the effect of gifted status on the compliers whose gifted status switches as their IQ score moves from just below to just above the Plan B threshold.

Specifically, our first-stage model takes the form:

$$Gifted_i = \pi \, 1[IQ_i \geq 116] + g(IQ_i) + u_i, \tag{2}$$

where $Gifted_i$ is an indicator that student i was identified as gifted by 5th-grade and $g(IQ_i)$ is a piece-wise linear function of IQ that allows a slope change at $IQ = 116$:

$$g(IQ_i) = g_0 + g_1(IQ_i - 116) + g_2(IQ_i - 116) \times 1[IQ_i \geq 116].$$

We estimate this model using data for students with IQ scores in some (possibly asymmetric) bandwidths (BW) around the 116-point threshold: $116 - BW_L \leq IQ_i \leq 116 + BW_R$.

Likewise, we estimate reduced-form models for the jump in an outcome of interest, Y_i, at the 116-point threshold of the form:

$$Y_i = \delta \, 1[IQ_i \geq 116] + h(IQ_i) + u_i, \tag{3}$$

where $h(IQ_i)$ is a piece-wise linear function analogous to equation (2). Our IV estimate of the treatment effect of gifted status on the mean outcome of the compliers is $\hat{\beta} \equiv \hat{\delta}/\hat{\pi}$. We construct this using a two-stage least squares (2SLS) estimator for the second-stage model with gifted status on the right-hand side:

$$Y_i = \beta \, Gifted_i + f(IQ_i) + e_i, \tag{4}$$

---

[22] This gap is similar in magnitude to the 15% gender gap in noncognitive skills that Cornwell et al. (2013) estimate using survey data from the 1998-99 Early Childhood Longitudinal Study.

using the indicator $1[IQ_i \geq 116]$ as an instrumental variable, and parameterizing $f(IQ_i)$ in the same way as $g(.)$ and $h(.)$. Our models have no covariates, but we show below that combinations of baseline characteristics that predict our outcomes are smooth at the 116-point threshold.

A standard approach to the choice of bandwidths BW_L and BW_R is the algorithm proposed by Calonico, Cattaneo and Titiunik (2014) (hereafter, CCT). Implicitly, this algorithm assumes that the running variable in the RD model is continuous. In our case, IQ is discrete with a relatively small number of values. We therefore proceed by considering results from a selection of bandwidths (though we also report results using the BW's suggested by CCT). In our baseline approach, we consider symmetric bandwidths of 6, 8 and 10 points to the left and right of the threshold, treating 8 as our preferred option. For most of our outcomes, the estimates are robust to the choice of bandwidth. For our main outcome, we present a more extensive analysis, including visualizations of the range of estimates from a full two-dimensional grid of alternative bandwidths BW_L and BW_R.

IV. Main Results

a. Validity of the RD Design

In an ideal RD design, the running variable is exogenously determined (Lee 2008; McCrary 2008). In the case of IQ tests, however, psychologists have some discretion in assigning scores and can boost marginal students above the gifted threshold.[23] Figures 2a and 2b show the frequency distributions of IQ scores for boys and girls in our analysis sample. Even though we use only first-reported IQ scores, the histograms show evidence of manipulation, with a spike in the frequency at 116 points and deficits at 114 and 115 points.[24]

Given this situation we use a parametric "donut" RD approach that excludes students with IQ scores of 114 to 116.[25] These specifications assume that the conditional means of the outcomes of interest are linear in IQ:

---

[23] For example, in one section of the Wechsler Intelligence Scale for Children (4th edition), children are given two words – like "red and blue" – and are scored 0, 1, or 2 points based on their explanation for their similarity.

[24] In a sample that uses our preferred IQ bandwidth and includes all IQ scores near the threshold for Plan B eligibility, we conduct a McCrary (2008) test and reject the null hypothesis of continuity at the threshold (p-value < 0.001). Using a donut approach that excludes IQ scores of 114 to 116, we fail to reject the null hypothesis of continuity at the threshold (p-value = 0.878).

[25] Another benefit of excluding IQ's just below the threshold is that it ensures our estimates are not confounded by psychological effects (e.g., disappointment) associated with just missing out on gifted status.

$$\begin{aligned} E[Y_i|IQ_i| &= h_0 + h_1(IQ_i - 116), & 116 - \mathrm{B}W_L \le IQ_i \le 116 \quad (5a) \\ &= \delta + h_0 + (h_1 + h_2)(IQ_i - 116), & 116 \le IQ_i \le 116 + \mathrm{B}W_R \quad (5b) \end{aligned}$$

allowing us to extrapolate from data outside the donut window back to the 116-point threshold and derive an estimate of δ. We make the same parametric assumption about our first-stage model to derive an estimate of π.

b. First-Stage Model and Compliers

Figures 3a and 3b show the relationships between IQ and the probability of being identified as gifted for boys and girls, respectively. We show the estimated probabilities for each IQ point (with the blue dots) along with pointwise 95% confidence intervals (indicated by the thin blue lines). The means for IQ scores inside the donut region are shown with open circles; we also shade this region. Finally, on each side of the donut, we show the estimated local linear fits for bandwidths of 6, 8, and 10 points, plotted in yellow, green, and red, respectively.

We note three features of the graphs. First, there are large jumps in the probability of gifted status at the Plan B threshold. To the left, the probabilities are low (around 8% for boys and 3% for girls, extrapolating to 116 points). To the right, the probabilities are much higher (around 58% for boys and 63% for girls). Second, gifted rates continue to rise with IQ to the right of the threshold. Third, the magnitude of the jump in the probability of gifted status is similar whether we use bandwidths of 6, 8, or 10 points for either boys or girls. Indeed, the estimates of our first stage model (presented in column 4 of Table 2) range from 47 to 52 ppts for boys and from 60 to 62 ppts for girls, with standard errors of around 5 ppts.

We use the method suggested by Abadie (2002) to estimate the characteristics of compliers whose gifted status is determined by whether their IQ's are marginally above or below the IQ threshold.[26] As shown in columns 5 and 6 of Table 1, compliers of both genders have relatively high 2nd and 3rd-grade scores: about 0.9σ's above the District-wide mean. This gap is about the same as the ~1σ gap between the Plan B IQ threshold and the population mean of IQ scores, suggesting that standardized test scores and IQ scores are generally aligned. Averaging test scores across subjects masks gender gaps by subject, however, with boys performing relatively well in math and girls in reading. There is also a large gender gap among compliers in

---

[26] Angrist, Hull and Walters (2023) describe how to apply this approach to estimate complier characteristics in a school lottery-based research design, and by extension, to an admissions RD, which they frame as a local lottery.

our non-cognitive measure, with 79% of girls but only 65% of boys classified as having high non-cognitive skills.

For a subset of cohorts, we predict college enrollment based only on 2nd-grade measures of cognitive ability (as in Figure 1) and estimate the complier means and quantiles of this index. The distribution of this index is very similar for boys and girls, with an overall mean of 0.78, and 10th and 90th percentiles of 0.72 and 0.81 respectively. These values are indicated with vertical lines on Figure 1 to illustrate that there are wide gender gaps in observed college entry rates, conditional on the index, in the range where most of our compliers are located.

c. The Effect of Gifted Identification on College Entry

We begin with the main outcome of our study: on-time completion of high school and entry to college in the following academic year (i.e., 7 years after completing 5th grade). Panel A of Figure 4 graphs mean rates of college entry by each level of IQ for boys and girls, using the same format as Figure 3. There is a clear jump in the college entry rate of boys at 116 points, although the precise magnitude depends on the bandwidth used to extrapolate from points outside the donut to the 116-point threshold, with a larger jump using a smaller bandwidth. In contrast, there is a relatively small jump (or possibly no jump) in the college entry rate for girls at the 116-threshold that is similar in size regardless of bandwidth.

As emphasized by Lee and Lemieux (2010), in a valid RD design, the predetermined characteristics of students with different IQ's should trend smoothly through the gifted threshold. We test this presumption in Panel B of Figure 4. Specifically, we use state-wide test scores and our non-cognitive index from 3rd grade, combined with student demographic variables, student cohort, and 5th-grade school characteristics, to predict the probability of college entry by gender.[27] We then plot the predicted probabilities at each IQ score, along with the associated confidence intervals. We also show estimated local linear models using bandwidths of 6, 8, and 10 points. Reassuringly, the predicted probabilities are very smooth through the 116-point threshold, and the estimated discontinuities are small.[28] Interestingly, the figure for boys also

---

[27] About 20% of gifted students are identified before 3rd grade, but there is no indication in our sample or in the previous literature that the services offered to gifted children in grades K-3 have any effect on test scores. We get very similar results if we replace 3rd-grade test scores with 2nd-grade SAT scores in math and reading or the 2nd-grade Nonverbal Ability Index from the test used for gifted screening (see Appendix Figure A3). But these measures are only available for about 70% of our sample, and are missing entirely for some cohorts.

[28] Appendix Table A4 reports estimates of discontinuities at the threshold for all the individual pre-determined student characteristics that we use to predict the probability of college entry.

shows evidence of deviations from the fitted lines for boys with IQ scores of 115 and 116, which is consistent with the type of selective manipulation that motivates our donut specification.

To probe the effects of bandwidth choice for our main outcome, we estimated a series of local linear specifications, using bandwidths of 3-14 points on left of the donut range (extending down to scores of 100) and 3-12 on the right (extending up to scores of 128).[29] We then constructed the full set of 120 estimates of the reduced-form jump in the probability of college entry. These are shown in Appendix Table A5 and plotted in Figure 5. The 2-dimensional surface of estimates for boys is relatively smooth and confirms that *smaller* bandwidths on either side of the threshold are associated with *larger* estimates of δ. Specifically, the estimated jump in college entry is around 30 ppt when both bandwidths are in the range of 3-4, but falls to around 7 ppt when the bandwidths are in the range of 12-14. For girls, the 2-dimensional surface is not as smooth, but the range of estimates of δ is smaller.

We also conduct two additional exercises to further assess our choice of bandwidth. First, we applied the CCT algorithm to our data, specifying an initial range of IQ's of 100-128 (deleting IQ's of 114-116), a rectangular kernel, and the default regularization procedure. The symmetric MSE-optimal bandwidth selected by the CCT algorithm is 3 for both boys and girls.[30] These choices lead to very large but imprecisely estimated reduced-form impacts of gifted eligibility for boys (32.0 ppt, s.e.=14.5 ppt), and small but imprecisely estimated impacts for girls (4.1 ppt, s.e.=14.1 ppt). Second, to evaluate the assumption of linear conditional expectations in equations (5a-5b) we constructed the root mean squared errors of models with different (symmetric) bandwidths (see Appendix Table A6). These are similar across bandwidths, as are the p-values of F-tests comparing linear models with given bandwidths to the unrestricted alternatives (with dummies for each IQ point). Given these results, and the visual evidence in Panel A of Figure 4, we believe that choices of 6, 8, and 10 are broadly representative of the data and that a bandwidth of 8 is a reasonable – if possibly conservative – choice for a preferred bandwidth.

---

[29] We stop at 128 points because there is a small spike in the rate of gifted identification at 130 points (the gifted threshold for non-disadvantaged students) and a deficit at 129 points. We believe these patterns are attributable to slippage in the applicability of the Plan B threshold, e.g., arising from changes in family incomes or ELL status.

[30] CCT's default triangular kernel selects a slightly wider symmetric bandwidth (4.8) for girls, but not for boys, and selects asymmetric bandwidths that are also similar for boys (3.7 below and 2.7 above) but slightly wider for girls (3.7 below, 5.4 above). Applying the CCT algorithm *without* regularization produces bandwidths that are generally larger: for boys this leads to a bandwidth of 3 points above the threshold and 10 below; for girls the choices are 7 points above the threshold and 12 below.

Table 2 summarizes the first-stage and reduced-form models for our main outcome and presents the corresponding 2SLS estimates of the effect of gifted identification. Panel A presents the results for boys and Panel B for girls. We show estimates, standard errors (clustering by school in 5th grade), and sample sizes for symmetric bandwidth choices of 8, 6 and 10 points; in all cases, we exclude IQ's between 114 and 116.

We begin in columns 1-3 by testing for discontinuities in student characteristics. The dependent variable in column 1 is a student's average 2nd-grade SAT score, which is only available for about 70% of our sample.[31] There is no indication of a discontinuity at the 116-point threshold. Column 2 presents models for 3rd-grade state-wide math and reading scores, which are available for a larger fraction of students in our sample: again, we find no evidence of any jump at the Plan B threshold. Finally, column 3 tests for jumps in the predicted probability of graduating on time and entering college, ass visualized in Panel B of Figure 4. These also show no evidence of discontinuities.

Next, in column 4, we present the estimated first-stage effect of crossing the 116-point threshold on the probability of being identified as gifted. As noted in the discussion of Figure 3, the estimates are centered around 0.5 for boys and 0.6 for girls, are robust to bandwidth choice, and are highly significant (t-statistic ≥ 10), addressing any concern about a weak first stage.

Column 5 presents the reduced-form discontinuities in the probability of graduating on time and entering college the next year. For boys, the estimates range from 0.09 to 0.19 and are all statistically significant, while for girls they are uniformly small (0.02-0.04) and insignificant. We have also estimated models for the probability of graduating high school on time or at most one year late, and for completing high school on time and entering college within two years (see Appendix Table A7). The effects on high school graduation are consistently small, suggesting that the main impact of gifted identification is on college entry, rather than on completing high school, and that the effects on college entry within 1 or 2 years are very similar.

Finally, column 6 presents 2SLS estimates of the treatment on the treated effects of gifted identification on college entry (i.e., β in equation 4). For boys, the point estimates range from 18 ppts for a 10-point bandwidth to 41 ppts for a 6-point bandwidth, with our preferred 8-point bandwidth yielding an estimate of 28 ppts (with a 95% confidence interval of 7-48 ppts). For girls, the estimates are narrowly clustered between 3 and 7 ppts, though the 95% confidence

---

[31] There is no detectable threshold crossing impact on the likelihood of having a 2nd-grade SAT score.

interval for our preferred 8-point bandwidth runs from -14 ppts to 24 ppts, so we cannot rule out a zero effect or a relatively large effect of either sign.

Figure 6 illustrates how these impacts affect gender gaps by presenting the estimated mean potential outcomes for the complying boys (left side) and complying girls (right side) using an 8-point bandwidth. The estimated college entry rate for male compliers who narrowly miss the gifted threshold is 46%; the 28 ppt treatment effect of passing the threshold raises this to 74%, which equals the rate for female compliers who narrowly miss the threshold. In other words, being identified as gifted raises the college entry rate of marginally eligible boys to the *same* rate as girls, closing the gender gap that is present in the absence of gifted services.

d. Interpreting Gender Differences in the Effects of Gifted Education

The estimates in Table 2 suggest that gifted identification raises the college entry rate of boys much more than that of girls, though we cannot quite reject that the IV estimates are the same at conventional significance levels: the p-value is 0.077 for our preferred 8-point bandwidth and slightly higher for the 6- and 10-point bandwidths. Given the visual evidence in Figures 4 and 5, however, we believe there is fairly strong evidence of a gender difference.

More importantly, the next section shows that there is a parallel set of effects on a variety of outcomes in middle and high school that are usually interpreted as markers of non-cognitive skill. Specifically, for outcomes like grades and advanced course enrollment, we find that boys perform worse than girls at IQ scores just below 116, and that passing the gifted threshold leads to gains for boys that close or nearly close the gap with girls, but has no significant effects for girls. In contrast, for outcomes like standardized test scores that are usually interpreted as measures of cognitive skill, we find small and insignificant effects for both boys and girls.

Building on the simple model in Section 1, our interpretation of these patterns is that in the absence of gifted services, many disadvantaged boys with IQ’s near the threshold have a non-cognitive skill deficit that limits their academic success. Most girls, on the other hand, have non-cognitive skills that are aligned with their cognitive skills. Assuming that non-cognitive and cognitive skills are strongly complementary, policies that boost non-cognitive skills but have no effect on cognitive skills will have positive effects for boys but little effect for girls.

This explanation suggests that if we were to compare girls and boys with similar levels of non-cognitive skills prior to receipt of gifted services, we would find that children with lower non-cognitive skills benefit from gifted services regardless of gender, while those with higher skills are less affected. To test this hypothesis, we stratify our sample using the indicator discussed above (and summarized in Table 1) for high 3rd-grade non-cognitive skills. We then fit RD models for the effect of gifted status on college entry separately for students with high vs. low non-cognitive skills, and we estimate the mean potential outcomes of the male and female compliers in each group. Our findings are summarized in Figure 7. Graphs of the corresponding first-stage and reduced-form models are presented in Appendix Figure A4.

Among students with lower non-cognitive skills – the majority of whom are male – the untreated compliers have relatively low college entry rates, whereas the treated compliers rates are closer to 80%. The implied effects of gifted services are large for both boys (58 ppt) and girls (27 ppt) but relatively imprecisely estimated. Among students with higher non-cognitive skills – the majority of whom are female – the untreated compliers have much higher college entry rates (60% for boys and 80% for girls) and the implied treatment effects of gifted services are small. There are still gender gaps in the treatment effect within each subgroup – which is not too surprising, given the limited power of our non-cognitive measure – but these gaps are much smaller than in the overall analysis sample, suggesting that differences by non-cognitive skill are the fundamental heterogeneity that drives the gender gap in the effect of gifted services.

e. Additional Subgroup Heterogeneity

Do the treatment effects of gifted services vary along dimensions other than gender and non-cognitive ability? We address this question in Appendix Figure A5, where we show 2SLS estimates of the effects of gifted identification from our preferred 8-point bandwidth models separately for different subgroups of boys and girls. We begin with race/ethnicity. For both boys and girls, the point estimates are slightly smaller for Black children, but not significantly so. Next, we separate students by socio-economic disadvantage and parental language. For girls, we see somewhat more positive estimates for those who do not receive free lunches and for those with non-English speaking parents, but we see no significant difference for boys. We also see no evidence of heterogeneity for either gender when we stratify by whether a student's average 3rd-grade test scores were above or below the median conditional on their IQ. Finally, stratifying students based on the share of FRL recipients or the average test scores at their schools again produces no evidence of heterogeneity for either boys or girls. In sum, we find no evidence of systematic heterogeneity by school characteristics or by student demographics other than gender, and no large differences within the gender groups other than the heterogeneity by non-cognitive skill noted in Figure 7.

V. Intermediate Outcomes

Next, we analyze a series of intermediate outcomes, starting in elementary school and proceeding through middle school and high school. We focus first on measures of advanced course selections and course grades – outcomes that are usually interpreted as markers of non-cognitive skill, especially when the level of cognitive skill is held constant (as it is in our RD models). We end the section by looking at standardized test scores, which we interpret as largely reflecting cognitive skills. To preview our results, we find that gifted identification leads to large positive treatment effects *only* for outcomes that reflect non-cognitive skill. Consistent with the model in Section I, most of these effects are positive and significant only for boys, whose outcomes in the absence of gifted services lag behind those of girls with similar cognitive ability.

a. Advanced Course Selections in Elementary, Middle and High School

In 2004 the District adopted a policy of assigning gifted children to separate classes in grades 4 and 5, comprised of gifted children or a combination of gifted and high-achieving students (see Section II.b). In column 1 of Table 3, we find large 2SLS estimated impacts of gifted status on the likelihood of enrolling in a GHA classroom in 4th grade, with impacts of 70-82 ppt for boys and 45-55 ppt for girls. To help visualize these effects, Figure 8 shows GHA participation rates of the compliers in our RD analysis with and without gifted status. Among the non-gifted compliers who narrowly missed the IQ threshold, there is a large gender gap in GHA participation, with rates of 26% for boys but 43% for girls. For gifted compliers (i.e., those who narrowly pass the IQ threshold) GHA participation rates are comparable and even slightly higher for boys than girls (97% versus 92%).

Further investigation into the participation gap among children with IQ's just below the gifted threshold reveals that only about one third it is due to higher scores among girls on the 3rd-grade statewide tests that are used to rank non-gifted students for assignment to a 4th-grade GHA class. The majority of the gap is due to different rates of participation between non-gifted girls and boys with similar test scores.[32] We suspect that this "preference for girls" may be due in part

---

[32] Consistent with results in Card and Giuliano (2016b, Figure 4) there is significant non-compliance in GHA enrollment both above and below the rank-based admissions threshold.

to teacher and parent concerns about placing boys with better test scores but lower non-cognitive skills in GHA classes.

The next three columns in Table 3 present estimated effects of gifted status on participation in the District's accelerated math curriculum ("GEM") and on algebra completion in middle school. Eligibility for GEM in 6th grade is based on 5th-grade math test scores. As shown in Appendix Figure D1, Panel B, there is a large (≈ 40 ppt) discontinuity in participation at the GEM test-score threshold among students in our IQ-based RD sample, very similar in size for boys and girls. Nevertheless, GEM participation rates among students with scores under the threshold are relatively high, and not all students who score above the threshold actually enter GEM. In this setting, eligibility for gifted services could impact GEM in two ways: by increasing 5th-grade math scores; or by increasing participation conditional on scores, either to the left of the cutoff (where schools may prioritize gifted students when filling extra seats) or to the right of the cutoff (where gifted students may receive extra encouragement to participate in GEM).

Column 2 investigates the first pathway, presenting 2SLS estimates of the effect of gifted identification on having an above-threshold math test score. The point estimates for boys are in the range of 8-12 ppt, but far from statistical significance (i.e., t-statistics ≤ 1); the estimates for girls are more variable across bandwidths but uniformly insignificant. Column 3 looks at the total effect on GEM entry, with a visualization of the associated reduced form in Panel A of Figure 9. The graphs suggest that boys with IQ's just under 116 are less likely to enter GEM than comparable girls, and that passing the gifted eligibility threshold is associated with a jump in participation for boys that largely closes the gap. In contrast, there is no apparent effect for girls. Consistent with these impressions, the 2SLS estimates in column 3 show a relatively large (and statistically significant) effect of gifted status on GEM entry for boys (on the order of 25 ppt), but weakly negative effects for girls.[33] We conclude that most of the rise in GEM participation for gifted boys is coming from an increase in enrollment conditional on test scores, rather than from a rise in the share of boys with scores above the GEM threshold.[34]

---

[33] Note that the reduced form jumps in Figure 8 are multiplied by a factor of about 2 for boys and 1.67 for girls to derive the 2SLS estimates.

[34] We have also examined the gender gap in GEM enrollment among non-gifted compliers (students with IQ scores just below the gifted cutoff). We find that the higher participation of girls is mainly driven by a higher rate of participation in GEM conditional on 5th grade math scores, rather than by higher test scores of girls in this IQ range. This finding, along with the similar finding about the gender gap in GHA participation described above, suggest that among students with IQ's just below the gifted threshold, gender gaps in course enrollments are caused by gaps in non-cognitive skills that makes parents and teachers less likely to push boys into advanced classes.

Column 4 looks at a related measure of middle school math participation: the probability of completing Algebra 1 by the end of 8th grade.[35] The reduced forms are visualized in Panel B of Figure 9. Again, the graphs show that boys with IQ's < 116 under-perform relative to girls, and that passing the gifted threshold leads to gains for boys that close the gap, while having little effect on girls. The estimated effects on completing algebra are a little larger than the effects on GEM, highly significant, and robust to bandwidth choices. Again, the effects for girls are all slightly negative though not significant.

The large magnitude of the impacts of gifted status on boys' participation in advanced math in middle school, coupled with the zero effects for girls, suggests that this could be an important channel mediating the effect on college entry. Indeed, other research (e.g., Card and Payne, 2020) suggests that success in math is especially important for the college entry of boys since a disproportionate share choose to major in math-intensive fields like business, computer science, or engineering. Using a supplemental research design, we directly estimate the magnitude of the impact of GEM on the rise in college entry rates for boys in Section V.e below.

Finally, column 5 of Table 3 and Panel C of Figure 9 examine a high-school outcome closely linked to college entry: the number of completed AP courses. The graphs show the same contrast between boys and girls with IQ<116 that we see for course selections in middle school: the number of AP classes is steadily rising in IQ for girls and reaches a level of about 4 courses for IQ's around 113-115 points, whereas for boys the relationship is much flatter and only reaches a level of about 2 courses. Crossing the 116-point threshold, however, pushes boys much closer to girls. Confirming the visual impression, the estimates in column 3 show a large effect of gifted status on boys' AP course-taking, with our preferred bandwidth yielding a +2.5 course effect (95% confidence interval = 0.7 to 4.2). In contrast, we see no effect for girls.

b. Course Grades and Disciplinary Actions

Table 4 reports results for additional non-cognitive outcomes in middle and high school, including course grades at both levels of schooling and rates of suspension in middle school. The estimates in column 1 show weakly positive point estimates for boys on middle school math grades but small negative estimates for girls. Given that gifted boys are more likely to take

---

[35] District middle schools offer an advanced math track in addition to GEM that provides access to Algebra-1 in 8th grade (rather than in 7th). Students who miss the GEM threshold (or score above the threshold but either do not enter or do not persist in GEM) may pursue the advanced track instead.

advanced courses, however, the fact that their grades are, if anything, slightly *higher* suggests some gain in achievement. In column 2, the dependent variable is an indicator for never having received a suspension (internal or external) in middle school. Here, we again see weakly positive estimated effects for boys and zero effect for girls. Though the estimates are all insignificant, plots of the reduced-form relationships, presented in Panel A of Figure 10, show patterns similar to those seen for the course-taking outcomes in Figure 9. There is a large gender gap at IQ's below the gifted threshold, with a "no-suspension" rate that is 15-20 ppt lower for boys than for girls. Crossing the threshold appears to narrow this gap.

The remaining columns in Table 4 examine high school grade point averages (on a 4-point scale) in math courses (column 3), English language arts courses (column 4), and all courses taken in high school (column 5). The reduced forms associated with columns 3 and 5 are plotted in Panels B and C of Figure 10. Both figures show that in the absence of gifted services, boys have much lower grades than same-IQ girls, with gaps of around 0.6 in math and 0.35 in overall high-school GPA. In the case of math grades, gifted status has marginally significant effects on boys that range from 0.37 to 0.50 – enough to reduce, but not quite close, the gap with girls. Interestingly, there is no corresponding effect for boys on language arts grades, and the net effect on overall GPA is small. And, in no case do we see significant effects for girls; the estimates are either close to zero or weakly negative.

c. Quality of Classroom Peers

Overall, we interpret the results in Table 3 and 4 as supporting the hypothesis that being identified as gifted increases boys' engagement in school, leading them to take more advanced math courses in middle school and earn higher math grades in high school. As noted, however, a concern for interpreting the effect of gifted status on course grades is that gifted students may take harder courses. In fact, there is evidence from responses to student questionnaires administered by the District that gifted students are guided toward more challenging classes. Column 1 of Table 5 presents 2SLS estimates of the effect of gifted status on responses by 6th graders to a question asking if they agreed with the statement: "This year, school staff helped me to select high level courses that challenge my abilities." (Graphs of the corresponding reduced-form relationships are presented in Appendix Figure A6).[36] The estimates suggest relatively large effects (around 0.7 σ's for boys and 1.3 σ's for girls).

To quantify the impacts of gifted status on the levels of classes actually taken by students, we use information on the early-grade test scores of classmates.[37] This approach sidesteps difficulties in distinguishing the level of a class from its title or number and provides a metric that can be compared across subject domains. Columns 2-7 of Table 5 summarize 2SLS estimates of the effect of gifted status on classmate test scores in middle and high school. Associated graphs of the reduced-form relationships are shown in Appendix Figure A7.

Inspection of the reduced-form graphs reveals patterns closely paralleling our findings for on-time college entry. For girls, there is a systematic positive relationship between IQ and peer quality throughout the full range of IQ's, with no evidence of a discontinuity around 116 points. In contrast, for boys with IQ's under the 116-point threshold, the relationship between IQ and peers is relatively flat, and at the gifted threshold we see jumps in peer quality that nearly close the gaps with similar-IQ girls.

Consistent with the reduced-form patterns, the 2SLS estimates for boys in Table 5 show positive effects of gifted identification on peer quality in all subjects and at all levels of schooling. The impacts are statistically significant for our preferred 8-point bandwidth (in the first row of Panel A) and centered around 0.3-0.4 standard deviation units. To the extent that teachers assign grades based on relative ranks, even the null effects on course grades in language arts for gifted boys noted in Table 4 might imply some improvement in effort and engagement.

In contrast to the effects of gifted status on peer quality for boys, for girls there is a significant effect only on middle school language arts classes. The differences between boys and girls are visualized in Figure 11, where we show the mean potential outcomes of peer quality for compliers in middle and high school.[38] At both levels, we find that untreated female compliers are in classes with higher scoring peers than untreated male compliers, and that the treatment of gifted status largely closes the gender gap.

---

[36] The responses are collected on a 5-point Likert scale; the standard deviation of responses is 1.24. The student questionnaire was only collected in the 2008-2011 academic years for 5th-grade cohorts so responses are missing for over one-half of our main sample. We find no statistically significant effect of gifted status on the probability of having a response (see Panel B of Appendix Figure A6).

[37] Specifically, we use state standardized test scores in reading and math, averaged over grades 3-5.

[38] See Appendix B, where we present bar charts similar to Figure 6 that compare potential outcomes of complying boys and girls, with and without gifted status, for all intermediate outcomes. The corresponding estimates for non-gifted compliers are reported in Appendix Table A3.

One possible explanation for the negligible effect on peer quality for girls is that they are already in the highest-level classes in their school. To test this, we re-estimated the reduced-form relationships shown in Appendix A7, deviating peer quality from the mean for all students in the same school. These graphs show very similar positive slopes between IQ and relative peer quality, even for girls with IQ's above the gifted threshold, suggesting that topping-out is not the explanation. Instead, we suspect that girls' class selections – like their other outcomes – are aligned with their cognitive skills, which as we show next were unaffected by gifted services.

The positive impact of gifted services on the peer quality experienced by gifted boys suggests that peer quality could one of the mediators that contributes to the increase in their college entry rates. Higher scoring peers could lead to a change in effort norms, or to a change in aspirations for post-secondary education. Unfortunately, it is very difficult to pin down the magnitude of such potential peer effects in our setting. We note, however, that two recent studies of the impact of selective high schools suggest that such peer effects are likely small.[39]

d. Cognitive Skills in Middle and High School

We conclude this section by examining whether gifted status affects *cognitive* skills as measured by test scores in middle and high school. Starting with middle school, Panels A and B of Figure 12 and columns 1-2 of Table 6 show results for standardized test scores in math and reading, averaged over grades 6-8. The graphs for both subjects show a strong positive relationship between IQ and test scores for both boys and girls, with roughly similar slopes for the two genders, and no evidence of a flatter relation for boys with IQ<116 (unlike the patterns seen in college enrollment and several of the non-cognitive outcomes).[40] Also, distinct from most of our "non-cognitive" outcomes for boys, Panels A and B show no evidence of discontinuity at the gifted threshold for either boys or girls – an impression that is confirmed by the small and insignificant treatment effect estimates in columns 1 and 2 of Table 6. Interestingly, while the reduced-form figures for math scores are similar for boys and girls, for reading, the entire test score-IQ relationship is lower for boys.

---

[39] Abdulkadiroglu et al. (2014) and Dobbie and Fryer (2014) both use regression discontinuity designs to study the effect of entry into selective high schools on college-related outcomes. Both studies find no impacts – see also Angrist (2014). In an earlier review of the peer literature, however, Sacerdote (2011) argues that peer effects in education tend to be positive.

[40] Converting IQ to standard deviation units, the slopes are around 0.7 – i.e., a 1σ rise in IQ's leads to about +0.7σ rise in standardized tests.

We reach similar conclusions about the effect of gifted status on cognitive skills when we measure these skills in high school. Here, we use the student's national percentile score on the PSAT exam, which is taken by most students in eleventh grade.[41] Figure 12 shows the reduced-form relationships between IQ and PSAT percentiles on the math (Panel C) and verbal (Panel D) subtests, while columns 3 and 4 of Table 5 present 2SLS estimates of the effect of gifted status. Looking first at the figures, we again see evidence that boys "under-perform" relative to girls on the verbal portion of the PSAT but not on the math portion. And again, there is no evidence of a discontinuity at the gifted threshold in boys' or girls' scores on either subject. These results are confirmed by the estimates in Table 5, which show no significant effect of gifted status on PSAT scores in either math or reading.

In sum, our analysis of intermediate outcomes suggests that gifted status pushes boys into more challenging courses and catches them up to girls with similar IQ's. The gendered effects on course selection are particularly striking in math, where boys just below the gifted threshold are less likely than their female counterparts to take advanced coursework, despite having similar test scores in math. Notably, gifted status also helps to narrow a wide gender gap in math grades despite having no detectable impact on test scores. Together, these results suggest that for boys near the 116 IQ threshold, gifted services "work" by narrowing the gap between their cognitive and non-cognitive skills, especially in the math domain.

VI. A Closer Look at Two Potential Mediators

In this section we take a closer look at two potential mediators of the effect of gifted identification on college entry: participation in GHA classrooms in grades 4 and 5, and participation in GEM in middle school. The analysis of GHA classes is complicated by the fact that under the District's policies, all gifted students are placed in these classes. An analysis of GEM is more straightforward because entry to the program is based on 5th-grade math test scores, and was not guaranteed for gifted students.

---

[41] The PSAT, short for Preliminary SAT, is administered during school as an opportunity for students to practice for the SAT; it also serves as the National Merit Scholarship Qualifying Test. We use the maximum score for students who take the PSAT more than once. We use PSAT instead of SAT scores because participation rates are much higher (around 92% vs. 70% for our cohorts), and also because the SAT scoring system was revised over our sample period whereas PSAT scores are reported to the District as national percentiles. Results for the SAT (available on request) are very similar but slightly noisier.

a. Gifted/High-Achiever Classrooms

Though nearly all gifted students are placed in GHA classes, non-gifted high achievers are also placed in these classrooms, leading to a fuzzy discontinuity in GHA participation at the gifted threshold. As seen in Figure 8, the size of the discontinuity is larger for boys than girls, largely because non-gifted girls with IQ's just below the 116-point threshold were more likely to be placed in a GHA class. This raises the question of whether some of the larger effect of gifted identification on boys' college entry can be attributed to their larger jump in GHA enrollment.

To address this question, we exploit variation across elementary schools in the numbers of gifted children per grade to stratify our sample into two groups: schools with relatively few GHA seats available for non-gifted high achievers, and those where more seats are available.[42] As we show in Appendix Figure C1, schools in the first group have less room for gender gaps in GHA participation for children without gifted status. As a result, the effect of gifted status on GHA placement is more similar between boys and girls at these schools (and if anything, is slightly larger for girls). However, there is no evidence of a treatment effect on college entry for girls in this sample, while the estimated effect for boys is even larger than in our overall sample. On the other hand, in the second group of schools, the treatment effect of gifted status on GHA participation is much smaller for girls, but the effects on college entry are more similar between boys and girls.[43] In light of these patterns, we find it unlikely that the small effect of gifted services on college entry for girls is due to their smaller average jump in GHA participation. Instead, we suspect that their relative abundance of non-cognitive skills explains the relatively small effects of gifted status for girls on *both* college entry and GHA participation.

A remaining question is whether GHA participation has gender-specific impacts on subsequent outcomes in middle and high school. Specifically, is GHA participation for boys the causal factor that leads to the non-cognitive gains found in our main analysis of the effects of gifted education? To study this question, we rely on an alternative sample that consists of *non-gifted* students whose 3rd-grade tests scores are near the top of their school/grade cohort. Our research design follows Card and Giuliano (2016b) and relies on the fact that open seats in a

---

[42] We also explored using variation in access to GHA classrooms based on cohort and age at gifted identification. However, the numbers of students who lacked access to a GHA classroom because the mandate was not in effect for their cohort or because they entered the gifted program too late are too small for meaningful sub-group analysis.

[43] Details on the construction of the two samples are in the notes to Appendix Figure C1. The corresponding estimates are presented in Appendix Table C1.

GHA classroom are allocated to non-gifted high-achievers based on relative ranks on an index of 3rd-grade scores. For comparability with our gifted analysis sample, we focus on boys and girls in this high-achiever sample who were FRL participants or English Language Learners. Unfortunately, the subset of non-gifted students whose GHA status in 4th and 5th grades is available in our data has very little overlap with the subset of students for whom we can observe high school outcomes and college entry. Thus, we limit our analysis to the impacts of GHA participation on middle school outcomes. The results are summarized in Appendix Table C1.

In brief, we find significant first-stage effects on GHA enrollment from having a 3rd-grade test score index above the presumed cutoff for a student's school and cohort. Reassuringly, we find no effect of test score rank on 3rd-grade test scores or on our measure of non-cognitive skill in 3rd grade (used in Figure 7), suggesting that the conditions for a valid RD are satisfied. Fitting 2SLS models for the various middle school outcomes, we find relatively large and marginally significant effects for boys on GEM enrollment, completion of algebra in middle school, the "no suspension" indicator, and math test scores. For girls, we generally find smaller and insignificant effects; however, the effect on combined test scores in reading and math (col. 8) is marginally significant and only slightly smaller than that for boys.

Whether these results translate to the students of primary interest in this paper – boys with IQ scores around the gifted threshold – is hard to know. Most of the compliers for gifted status are <u>not</u> high achievers (as evidenced by the low rate of GHA participation for non-gifted boys, noted in Figure 8). Moreover, as shown in Appendix Table A1, the two groups differ on several other dimensions. Perhaps most importantly, the estimates in Appendix Table C2 suggest (consistent with the findings of Card and Giuliano 2016b) that for non-gifted high achievers of both genders, GHA entry leads to sizeable gains in standardized test scores. In contrast, we find no impact of the entire package of gifted services (including GHA classrooms) on the test scores of boys or girls with IQ's around the gifted threshold. Thus, while it is possible that the effects of GHA classes on the *non-cognitive* skills of compliers at the gifted threshold are similar to the effects we can measure for high achievers, the differential impacts on *cognitive* skills between these two groups suggests that we should be cautious about extrapolating.

b. GEM

Gifted status is associated with a rise in GEM participation for boys, along with subsequent gains in various math-related outcomes in high school. The fact that GEM <u>precedes</u>

the changes in high school course-taking and achievement raises the question of whether GEM is a significant mediator of the rise in college entry for gifted boys. To answer this, we estimate the causal effects of GEM entry directly, using the 5th-grade math test score cutoff for GEM eligibility. Using the subset of students in our main analysis sample whose 5th-grade math test scores lie within a narrow bandwidth of the GEM threshold, we conduct an RD analysis of the effects of entering GEM in 6th grade on the probability of graduating high school on time and entering college (i.e., our main outcome). If GEM participation is in fact the key driver of the gains in college entry for gifted boys, then we expect these RD models to show large causal effects – large enough to account for all or most of the effect of gifted status on college entry.

The results are presented in Appendix Table D1, which closely parallels Table 2. To summarize, we find that pre-determined student characteristics trend smoothly through the GEM cutoff, but GEM participation rises sharply at the threshold, with jumps for both boys and girls in the range of 35 to 42 ppt that are robust to the choice of bandwidth.[44] We find a positive and statistically significant reduced-form effect of passing the GEM threshold on college entry for boys, on the order of 11 ppts, and a 2SLS estimate of the effect of GEM of 28 ppts that is marginally significant (*t*–statistic ≈2). By comparison, the reduced-form and 2SLS estimates for girls are all relatively small in magnitude and insignificantly different from zero.[45]

Taken at face value, the estimated 28 ppt effect of GEM on college entry for boys, coupled with the fact that gifted status leads to 26 ppt increase in GEM participation of boys (Table 3) implies that GEM can account for only about <u>one-quarter</u> of the overall 28 ppt effect of gifted status on college entry of boys. Moreover, as shown in Appendix Figure C4, there is no indication that GEM participation has any effect on high school math grades or AP courses of boys. Thus, we conclude that GEM participation is probably <u>one</u> of the channels contributing to the effect of gifted status on boys' college entry rates, but that gifted identification has other effects, including impacts on high school course selection, that are not mediated by GEM.[46]

---

[44] Appendix Figure D1 illustrates the reduced-form relationship between the running variable and GEM participation (Panel B) and the analogous relationships for gifted status (Panel C).

[45] Appendix Figure D2 shows the reduced-form relationships for boys and girls using the three selected bandwidths, and Appendix Figure D3 plots reduced-form estimates for the full range of bandwidths between 5 and 30.

[46] A caveat – suggesting that some caution is warranted when extrapolating from the RD results in Appendix D to the gifted population – is that the compliers with the 5th-grade math score cutoff for GEM differ from the gifted compliers who are pushed into GEM when their IQ score exceeds 116. As shown in Appendix Table D2, IQ-based compliers have lower average test scores and lower non-cog skills than compliers based on 5th-grade math scores.

VII. Interpretation and Conclusions

Students who are identified as gifted in the District receive a package of services, including some individualized instruction in grades 1-3; assignment to separate gifted/high achiever classrooms in 4th and 5th grades; and biannual meetings with gifted specialists throughout their school careers. Despite the modest nature of these services, they appear to have relatively large effects on the college entry rate of boys, and on their middle and high school course selections and grades. At the same time, there is little evidence of corresponding effects for girls, or on standardized test scores of either gender.

Table 7 summarizes our main findings, grouping outcomes by whether they are usually interpreted as markers of non-cognitive skill (Panel A), cognitive skill (Panel B) or longer-run achievement (Panel C), and sorting within each group by the level of schooling. Column 1 of the table shows the mean potential outcomes for non-gifted female compliers. Column 2 shows the gender gap in mean outcomes of the non-gifted compliers. Column 3 shows the estimated treatment effect of gifted services on the outcome for boys, and column 4 shows the gap between the treatment effects for boys and girls.

As we have seen in many of the previous figures, the results in columns 1 and 2 show that boys with IQ scores just under 116 points appear to have substantially lower non-cognitive skills than similar-IQ girls, with gender gaps of around 30-50% of the girls' mean for many outcomes in Panel A (e.g., a 47% gap in GEM enrollment and a 46% gap in the number of AP courses completed). The gender gaps in measures of cognitive skills (Panel B) are generally smaller, particularly in the math domain (e.g., only a 0.04σ gap in grade 6-8 math scores, and essentially a zero gap in PSAT math scores). Finally, the gender gap in the college entry rate of the non-gifted compliers is also large (0.28 for boys versus 0.74 for girls, equivalent to 38% of the girls' mean).

We see positive treatment effects of gifted services on nearly all the non-cognitive outcomes for boys, coupled in most cases with small effects on girls, so the relative treatment effect on boys versus girls (column 4) is often about as large as the boys' effect (in column 3).[47]

---

[47] To address concerns about multiple hypothesis testing, we report a parallel set of treatment effect estimates in Appendix Table A8, along with the conventional p-values and adjusted "*q*-values" that control for the false discovery rate (FDR) using the two-step procedure from Benjamini, Krieger, and Yekutieli (2006). All of the estimates that are at least marginally significant using conventional p-values remains so, and the effect on college enrollment remains significant with an adjusted q-value of .022.

Interestingly, many of the relative treatment effects on boys are also large enough to fully offset the gender gap in means of the untreated compliers (i.e., the entry in column 4 is about the same size as the entry in column 2), meaning that gifted services appear to close or nearly close the gender gap in non-cognitive skills. In contrast, none of the treatment effects on cognitive outcomes of boys (or girls) is even marginally significant.

How can we interpret this combination of effects? Previous research (e.g., Heckman et al., 2006) has argued that cognitive and non-cognitive skills are strong complements in the production of educational attainment. Other work (e.g., Jacob, 2002; Cornwell et al., 2013) has asserted that boys have lower non-cognitive skills than girls with similar cognitive skills. Putting these two strands of work together, we showed in Section I that an intervention that raised non-cognitive skills, with no effect on cognitive skills, would be expected to improve the educational outcomes of boys, while having small impacts on most girls. Differentiating between skills in the math and non-math domains provides an even richer set of implications, since boys who are marginally qualified for gifted identification tend to have relatively high cognitive skills in math. In that case we might expect an intervention that raised non-cognitive skills to yield particularly large gains in math-related course selections and grades for boys in middle and high school.

We believe this very simple framework provides a successful interpretation for our main results. The gifted education literature has long emphasized the importance of targeting gifted services to help children with high cognitive abilities achieve their full potential. A model with strong complementarity between cognitive and non-cognitive skills provides a plausible framework for interpreting this goal. We acknowledge that some researchers (e.g., Lundberg, 2020) have argued that the gender gap in schooling outcomes arises not from a non-cognitive skills deficit but because boys have gender-stereotypical aspirations that lead to negative attitudes toward school. This hypothesis has many of the same implications as one based on non-cognitive skills. Further research could benefit from settings with richer measures of non-cognitive skills that could be compared across age groups to verify that these skills are raised by gifted services, and potentially to test whether non-cognitive skills vary across subject domains.

Overall, our findings add to the growing evidence that long-run schooling outcomes depend on more than just cognitive ability; that the other determinants of school success (non-cognitive skill, aspirations, engagement) are differentially distributed between boys and girls; and that "small" interventions can substantially affect schooling outcomes via these other channels. Importantly, even relatively large impacts on these alternative channels may not be detected by standardized tests. In our case, gifted status has no significant effect on standardized tests measured in grades 5-8, even though it causes large changes in boys' middle school math curriculum and grades, and in their classroom peers. Moreover, there is no effect on PSAT scores, despite boosts in high school math GPA's, high school classroom peer quality, and the number of AP classes.

With respect to the gifted education literature, our findings provide the first rigorous evidence from a U.S. setting that gifted programming can have long term effects on disadvantaged students. Previous evaluations (e.g., Bui et al, 2014; Card and Giuliano, 2014) have focused on standardized test scores and found no impact. Our findings, and related findings by Cohodes (2020), suggest that it is important to look beyond standardized test scores to understand the value of advanced academic programming. Since disadvantaged boys (even those with high cognitive abilities) have low college entry rates, this is a natural metric in our setting. For non-disadvantaged student populations, however, it may be necessary to look beyond college entry and focus on outcomes like advanced degree attainment, field of study, or even earnings (Booij et al., 2016; Lavy and Goldstein, 2022).

References

Abadie, Alberto. 2002. “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models.” *Journal of the American Statistical Association* 97(457):284–92.

Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak. 2014. “The Elite Illusion: Achievement Effects at Boston and New York Exam Schools.” *Econometrica* 82 (1): 137–96.

Angrist, Joshua D. 2014. “The perils of Peer Effects." *Labour Economics* 30(C): 98-108.

Angrist, Joshua D., Peter Hull, and Christopher Walters. 2023. “Methods for Measuring School Effectiveness” In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann (eds.) Handbook of the Economics of Education, Vol. 7. Elsevier.

Aucejo, Esteban M., and Jonathan James. 2019. “Catching Up to Girls: Understanding the Gender Imbalance in Educational Attainment within Race.” *Journal of Applied Econometrics* 34(4): 502-525.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. “Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes.” *American Economic Journal: Applied Economics* 11(3): 338–381

Barreca, Allan I., Jason M. Lindo, and Glen R. Waddell. 2011. “Heaping-Induced Bias in Regression-Discontinuity Designs.” National Bureau of Economic Research (NBER) Working Paper 17408.

Becker, Gary S., William H.J. Hubbard, and Kevin M. Murphy. 2010. “Explaining the Worldwide Boom in Higher Education of Women.” *Journal of Human Capital* 4 (3): 203–41.

Benjamini, Y., A.M., Krieger and D. Yekutieli. 2006. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*. 93(3): 491-507.

Bertrand, Marianne, and Jessica Pan. 2013. “The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior.” *American Economic Journal: Applied Economics* 5 (1): 32–64.

Bhatt, Rachana. 2010. “The Impacts of Gifted and Talented Education.” Unpublished Working paper, Georgia State University.

Betts, Julian R. 2011. The Economics of Tracking in Education. In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*. Volume 3. New York: Elsevier. Pages 341-381.

Booij, Adam, Ferry Haan, and Erik Plug. 2016. “Enriching Students Pays Off: Evidence from an Individualized Gifted and Talented Program in Secondary Education.” IZA Discussion Paper No. 9757.

Bowles, Samuel. 1970. "Towards an Educational Production Function," NBER Chapters, in: *Education, Income, and Human Capital*, pp. 11-70, National Bureau of Economic Research.

Bui, Sa A., Steven G. Craig, and Scott A. Imberman. 2014. "Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Students." *American Economic Journal: Economic Policy* 6(3): 30-62.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica* 82(6): 2295-2326.
Card, David and Laura Giuliano. 2014. "Does Gifted Education Work? For Which Students?" National Bureau of Economic Research Working Paper Number 20453.

Canaan, Serena, Stefanie Fischer, Pierre Mouganie, and Geoffrey C. Schnorr. 2024. "Keep Me in Coach: The Short- and Long-Term Effects of a University Coaching Intervention." Unpublished Manuscript.

Card, David and A. Abigail Payne. 2021. "High School Choices And The Gender Gap In Stem," *Economic Inquiry* 59(1): 9-28.

Card, David and Laura Giuliano. 2016a. "Universal Screening Increases the Representation of Low-income and Minority Students in Gifted Education." *Proceedings of the National Academy of Sciences* 113(48): 13678-13683.

Card, David and Laura Giuliano. 2016b. "Can Tracking Raise the Test Scores of High- Ability Minority Students?" *American Economic Review* 106(10): 1-35.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti. 2022. "Goals and Gaps: Educational Careers of Immigrant Children." *Econometrica* 90(1): 1-29.

Charles, Kerwin Kofi and Ming-Ching Luoh. 2003. "Gender Differences In Completed Schooling." *Review of Economics and Statistics* 85(3): 559-577.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi. 2016. "Childhood Environment and Gender Gaps in Adulthood." *American Economic Review*106 (5): 282–88.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017. "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." National Bureau of Economic Research Working Paper Number 23618.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *Quarterly Journal of Economics* 135(2): 711-783.

Cohodes, Sarah R. 2020. "The Long-Run Impacts of Specialized Programming for High-Achieving Students." *American Economic Journal: Economic Policy* 12(1): 127-166.

Cornwell, Christopher, David B. Mustard, and Jessica Van Parys. 2013. "Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary School." *Journal of Human Resources* 48(1): 236-264.

Cunha, Flavio and James J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97(2): 31-47.

Dobbie, Will, and Roland G. Fryer Jr. 2014. "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools." *American Economic Journal: Applied Economics* 6 (3): 58–75.

Florida Department of Education. 2006. "Development of Educational Plans for Students Who Are Gifted." Technical Assistance Paper, Paper Number FY 2006-3.

Ford, Donna Y. 1998. "The Underrepresentation of Minority Students in Gifted Education: Problems and Promises in Recruitment and Retention." *Journal of Special Education* 32(1): 4-14.

Fortin, Nicole M., Philip Oreopoulos, and Shelley Phipps. 2015. "Leaving Boys Behind: Gender Disparities in High Academic Achievement." *Journal of Human Resources* 50 (3): 549–79.

Goldin, Claudia, Lawrence F. Katz and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20(4): 133–156.

Hahn, J., Todd, P., & Van der Klaauw, W. 2001. Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica* 69(1): 201–209.

Hanushek, Eric A. 2020. "Education Production Functions." In Steven Bradley and Colin Green (Eds.) *The Economics of Education: A Comprehensive Overview*, pp. 161-170.

Heckman James J. and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91(2): 145-149.

Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312 (30): 1900-1902.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3): 411–82.

Imbens, Guido and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467-475.

Jackson, C. Kirabo. 2018. “What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes.” *Journal of Political Economy* 126(5): 2072-2106.

Jacob, Brian A. 2002. “Where the Boys Aren’t: Non-cognitive Skills, Returns to School and the Gender Gap in Higher Education.” *Economics of Education Review* 21: 589-598.

Lavy, Victor and Yoav Goldstein. 2022. “Gifted Children Programs’ Short and Long Term Impact: Higher Education, Earnings, and the Knowledge Economy.” NBER Working Paper Number 29779.

Lee, David S. 2008. “Randomized Experiments from Non-random Selection in U.S. House Elections.” *Journal of Econometrics* 142(2):675–97.

Lee, David S., and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature* 48 (2): 281–35.

Lundberg, Shelly. 2020. “Educational Gender Gaps.” *Southern Economics Journal* 87(2): 416-439.

McCrary J. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics* 142(2):698–714.

Murphy, Patrick R. 2009. “Essays on Gifted Education's Impact on Student Achievement.” Unpublished Doctoral Dissertation, Department of Economics, Florida State University.

Rampino, Tina, and Mark P. Taylor. 2013. “Gender Differences in Educational Aspirations and Attitudes.” ISER Working Paper Series, No. 2013-15.

Rose, Evan K., Jonathan T. Schellenberg and Yotam Shem-Tov. 2002. “The Effects of Teacher Quality on Adult Criminal Justice Contact.” NBER Working Paper 30274.

Sacerdote, Bruce. 2011. “Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?” In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*. Volume 3. New York: Elsevier. Pages 249-277.

Subotnik, Rena F., Paula Olszewski-Kubilius, and Frank C. Worrell. 2011. “Rethinking Giftedness and Gifted Education: A Proposed Direction Forward Based on Psychological Science.” *Psychological Science in the Public Interest* 12(1): 3-54.

Todd, Petra E. and Kenneth I. Wolpin. 2003. “On The Specification and Estimation of The Production Function for Cognitive Achievement.” *Economic Journal* 113(485): 3-33.

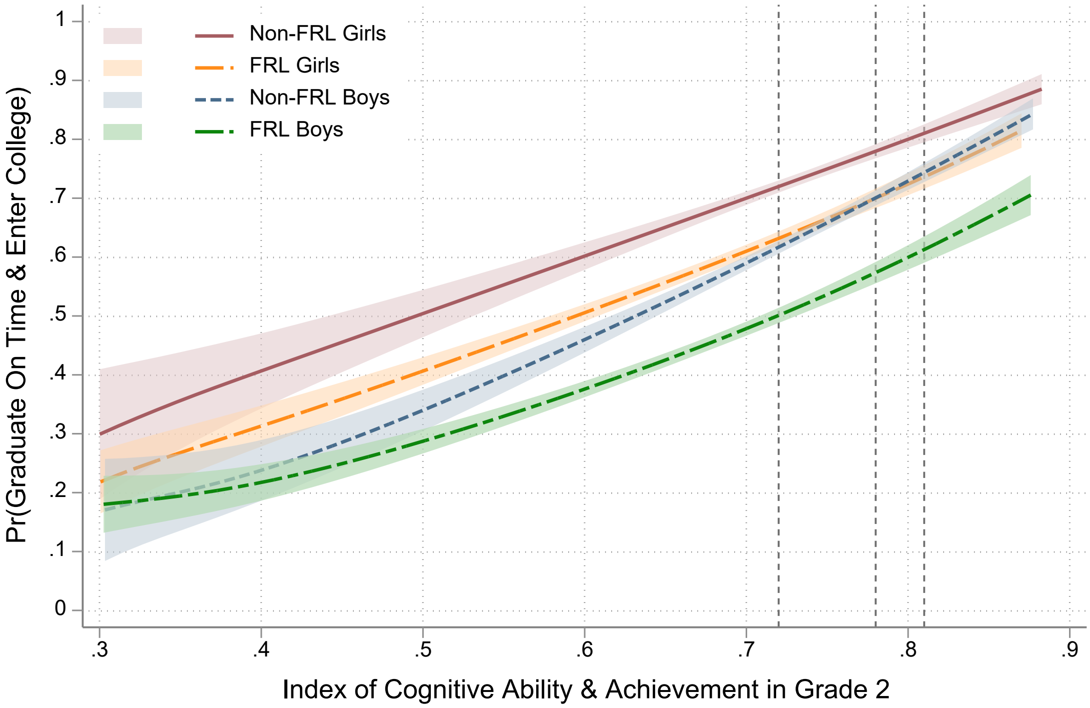
VanTassel-Baska, Joyce. 2005. “Gifted Programs and Services: What Are the Nonnegotiables?” *Theory Into Practice* 44(2): 90-97.

VanTassel-Baska, Joyce. 2018. “American Policy in Gifted Education.” *Gifted Child Today* 41 (2): 98-103.

Vincent-Lancrin, Stephan. 2008. “The Reversal of Gender Inequalities in Higher Education: An On-going Trend.” In *Higher Education to 2030.* Volume 1: Demography. Paris: OECD.

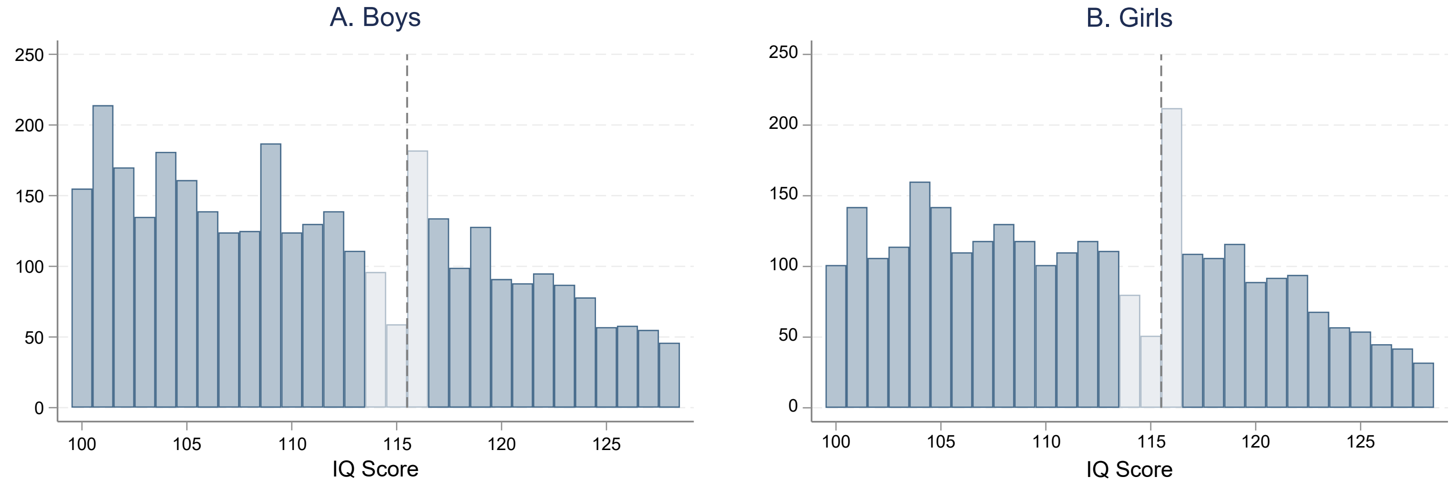
Zirkel, Perry A. 2005. “State Laws for Gifted Education: An Overview of the Legislation and Regulations.” *Roeper Review* 27(4):28-232.

Figure 1. Gaps in College Entry by Cognitive Ability & Achievement in Grade 2



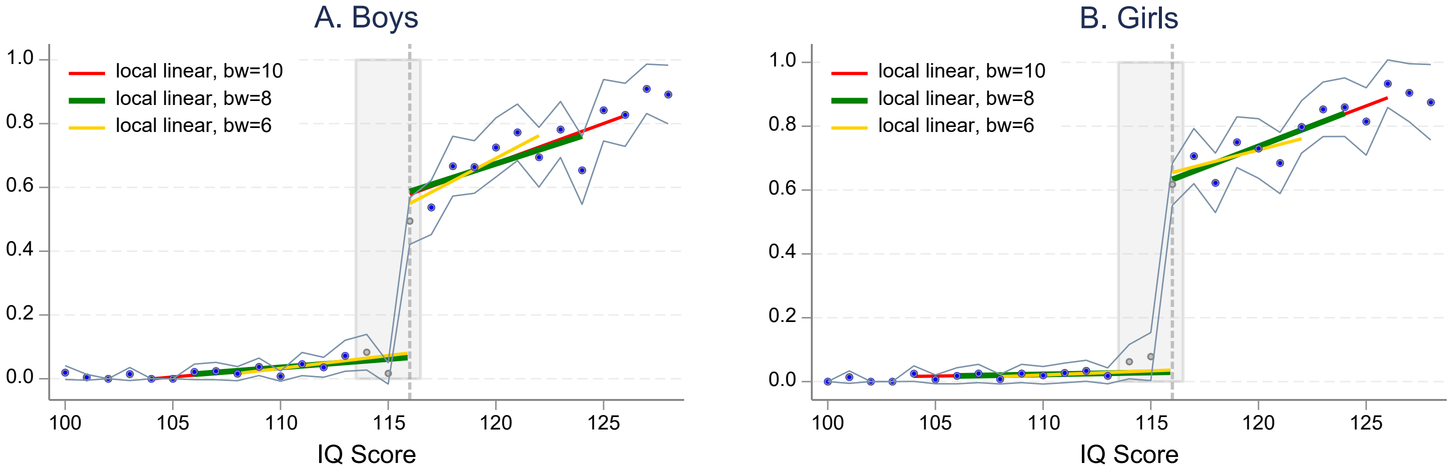
Notes: This figure plots estimated means and 95% confidence intervals of an indicator for graduating high school on time and enrolling in college against an index of cognitive ability and academic achievement in grade 2. The index is constructed as the predicted probability of graduating on time and entering college from a model fit to all non-gifted, non-FRL girls in the sample; hence, the means for Non-FRL Girls (represented by the maroon line in the figure) fall along the 45% line in the figure by construction. The means for the other three groups – FRL Girls (orange), Non-FRL Boys (blue), and FRL Boys (green) – at each value of the index can be interpreted as the rate of on-time graduation and college entry for students in that group with the same cognitive ability and second-grade achievement as the Non-FRL girls whose rate is given by the index. Ability and achievement are measured using scores from three tests administered to all second graders in the studied district in the spring of 2005, 2006, and 2007: the Naglieri Nonverbal Ability Test (NNAT) and the Stanford Achievement Tests (SAT) in math and reading. The sample includes all students in these three cohorts who were not identified as gifted by the end of fifth grade and who could be followed in our data for at least 10 years after they were enrolled in grade 2. The means and confidence intervals are smoothed using a local linear bi-weight kernel regression with a bandwidth of 0.45. The dashed vertical line at 0.78 represents the mean value of the index among the compliers in our IQ-based regression discontinuity analysis for the effects of gifted status. The dashed vertical lines at 0.72 and 0.81 represent the 10th and 90th percentiles of the index among sample compliers.

Figure 2. Histograms of Running Variable (First IQ Score), by Student Gender



Notes: This figure shows frequency distributions of the student IQ scores used to estimate our RD models separately for boys (Panel A) and girls (Panel B). For students with multiple IQ scores on record, we use the score from the first time the student was evaluated. The pooled sample consists of all students in the District who entered grade 5 in 2003-2012; who had an IQ score by the end of grade 5; who were eligible for a free or reduced-price lunch (FRL) and/or were enrolled in the English-language learner (ELL) program at the time of their IQ test; and whose first recorded IQ score is 100-128. The dashed vertical line indicates the threshold of IQ≥116 for gifted status under the Plan B eligibility criteria for FRL/ELL students. The range of scores shaded in light blue (114-116) is the "donut" hole that we exclude from the estimation sample to mitigate concern about selective manipulation of IQ scores.

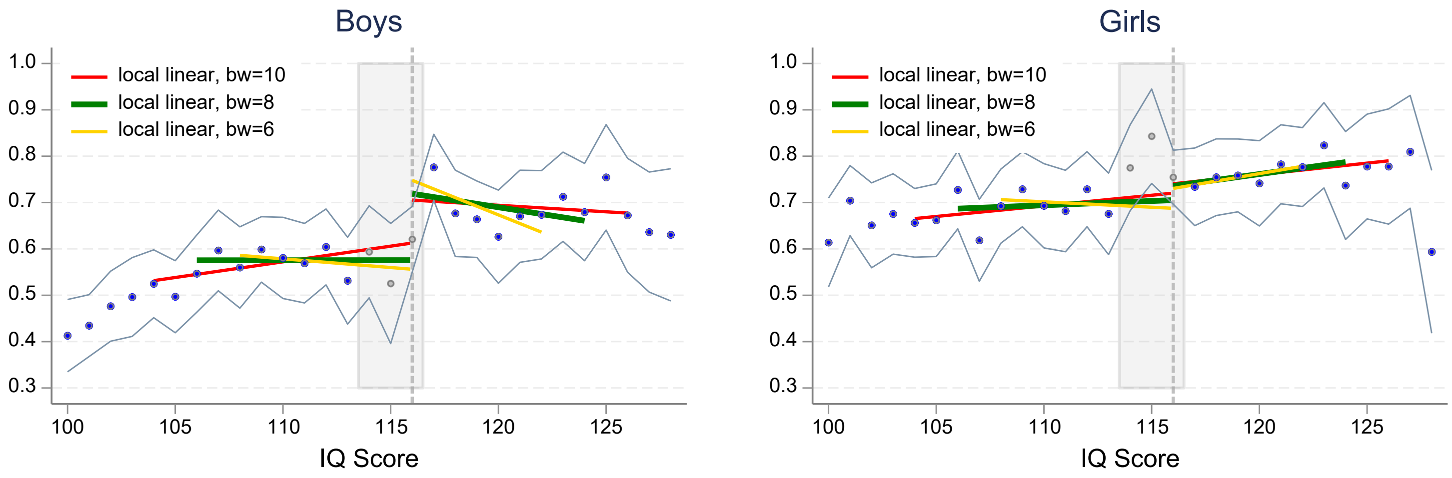
Figure 3. First-Stage Relationship for Gifted Status



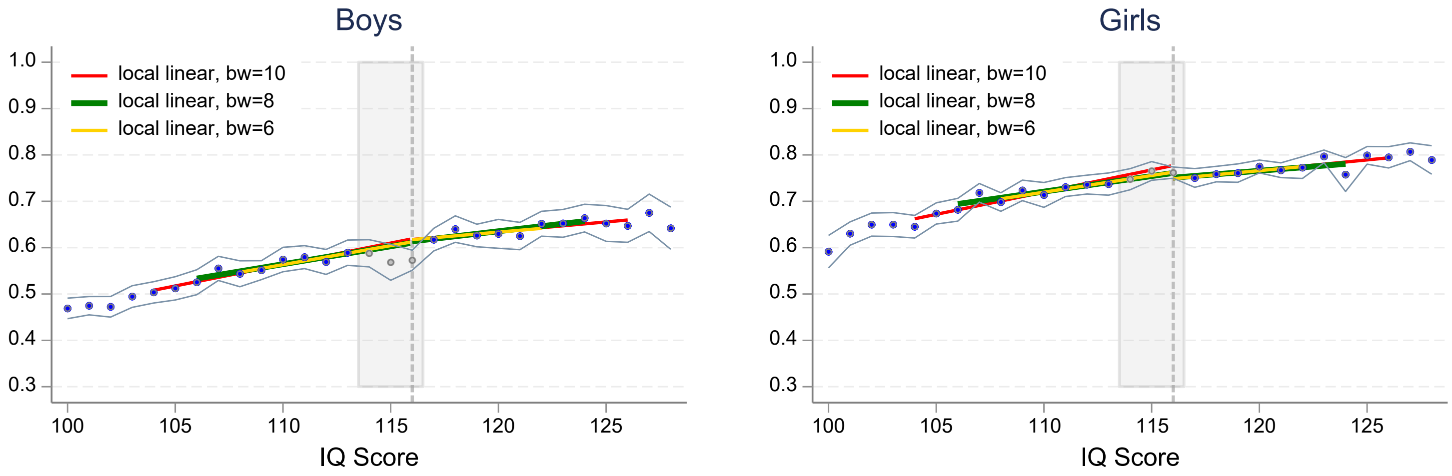
Notes: This figure plots sample means (blue dots) of an indicator for gifted status (as of the end of 5th grade) at each value of IQ, for boys (Panel A) and girls (Panel B). It also shows 95% confidence intervals for the IQ-specific means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles.

Figure 4: Reduced-Form Relationships for On-Time Graduation and College Enrollment

A. Observed

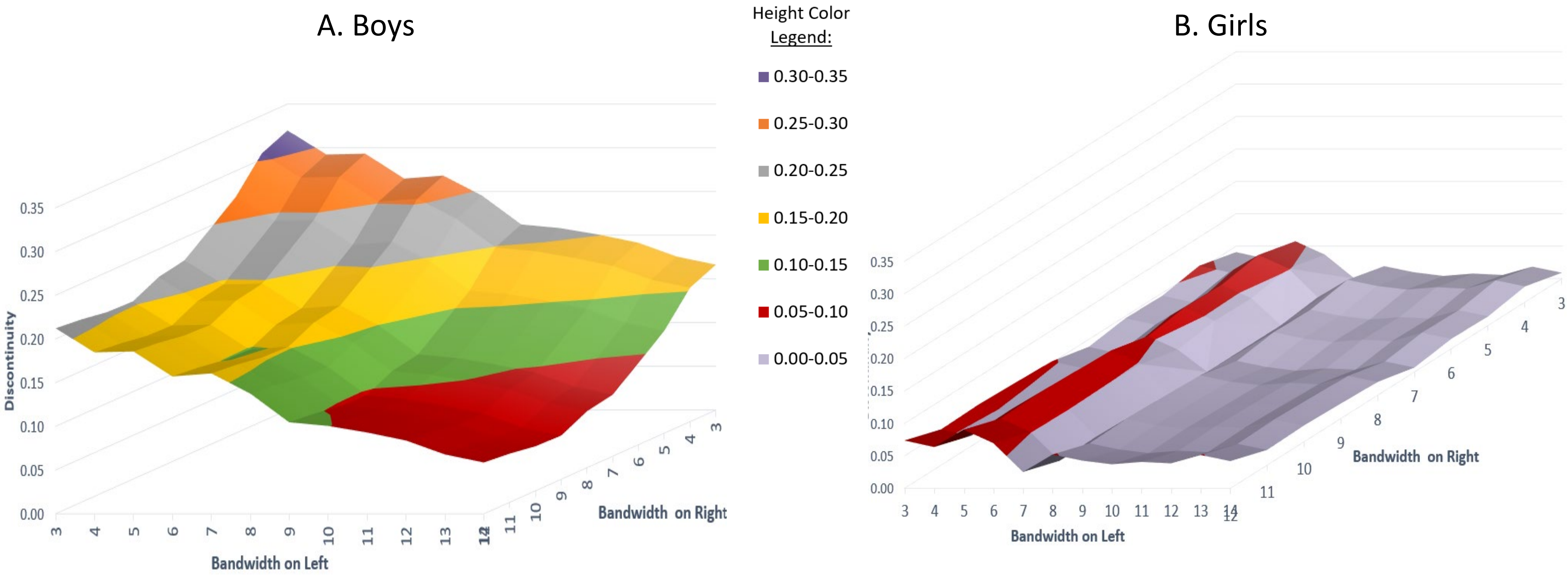


B. Predicted



Notes: This figure plots sample means (blue dots) of the dependent variable at each value of IQ, for boys (left) and girls (right), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. In Panel A (top row), the dependent variable is an indicator for graduating high school on time and enrolling in college within one year. In Panel B (bottom row), the dependent variable is the predicted probability of graduating on time and enrolling in college from models fit separately to boys and girls with IQ scores in the range from 100 to 115. The prediction models include second-order polynomials in 3rd-grade math and reading test scores; students' responses in grade 3 to a survey question about how much they enjoy learning in their school; the number of internal and external suspensions in grade 3; indicators for student race/ethnicity, FRL status, ELL status, and cohort; average test scores, fraction FRL and fraction nonwhite of the school where the student is enrolling in grade 5; and the median household income of the student's neighborhood.

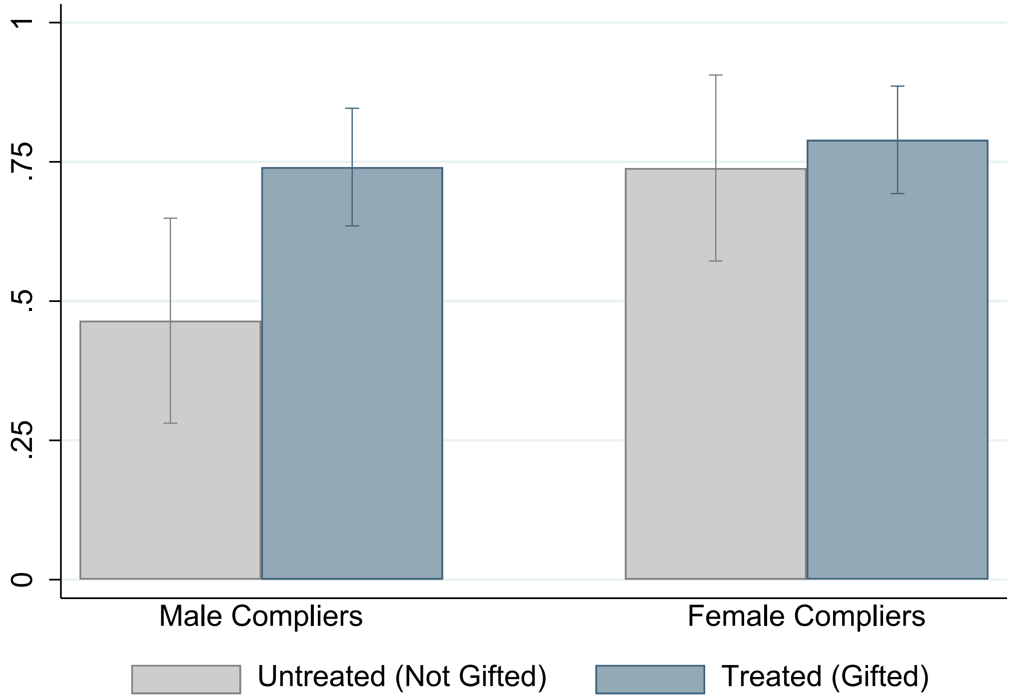
Figure 5: Estimated Discontinuities with Varying Bandwidths on Left and Right: Graduate High School On Time and Enter College


A. Boys

B. Girls

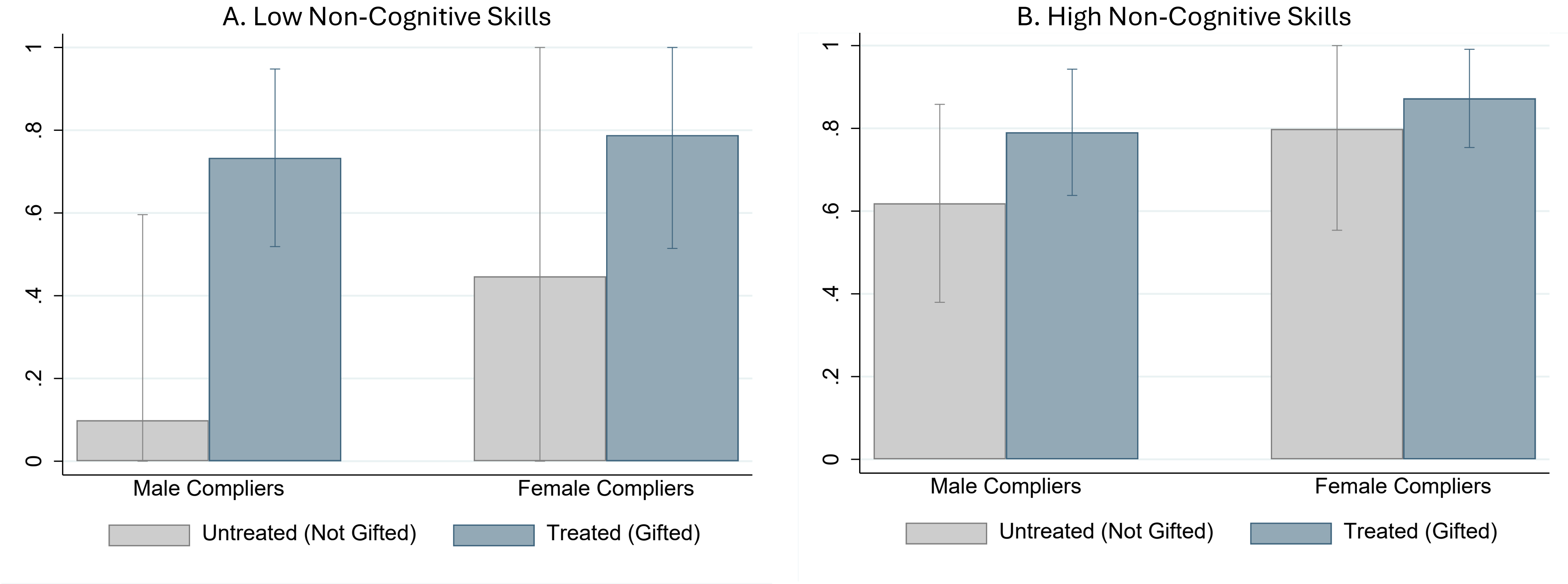
Notes: This figure plots reduced-form estimates from local linear RD models with varying bandwidths (IQ points) to the left and right of the gifted eligibility threshold, separately for boys (Panel A) and girls (Panel B). The dependent variable is an indicator for graduating from high school on time and enrolling in college the following year. All models exclude IQ scores that fall in the "donut" region (114-116).

Figure 6: On-Time Graduation & College Enrollment, Gifted and Non-Gifted Compliers



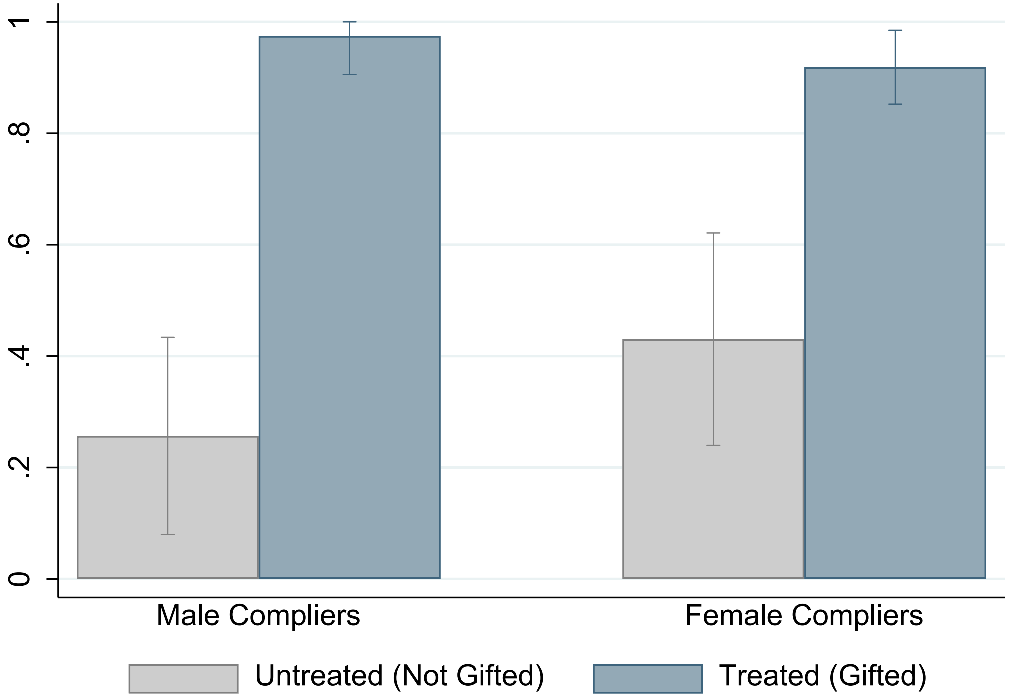
Notes: This figure shows the estimated probabilities of graduating high school on time and enrolling in college within one year for gifted and non-gifted compliers. The estimates are constructed following the approach of Abadie (2002). Specifically, complier untreated outcomes are calculated using a linear RD IV specification where the dependent variable is specified as the interaction between on-time graduation and college enrollment and an indicator for not being gifted. Complier treated outcomes are calculated similarly where the dependent variable is an interaction between the outcome and an indicator for gifted status. All estimates are constructed using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the “donut” region (i.e., scores 114-116).

Figure 7: On-Time Graduation & College Enrollment, Gifted and Non-Gifted Compliers By Gender and Non-Cognitive Skill Index in Grade 3



Notes: This figure shows the estimated probabilities of graduating high school on time and enrolling in college within one year for gifted and non-gifted compliers, estimated separately for students with low (Panel A) vs. high (Panel B) non-cognitive skills in grade 3. Students are classified as having high non-cognitive skills if they had no disciplinary actions in 3rd grade and if, in response to a District-administered survey, they "strongly agreed" that they enjoyed learning at their school. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.)

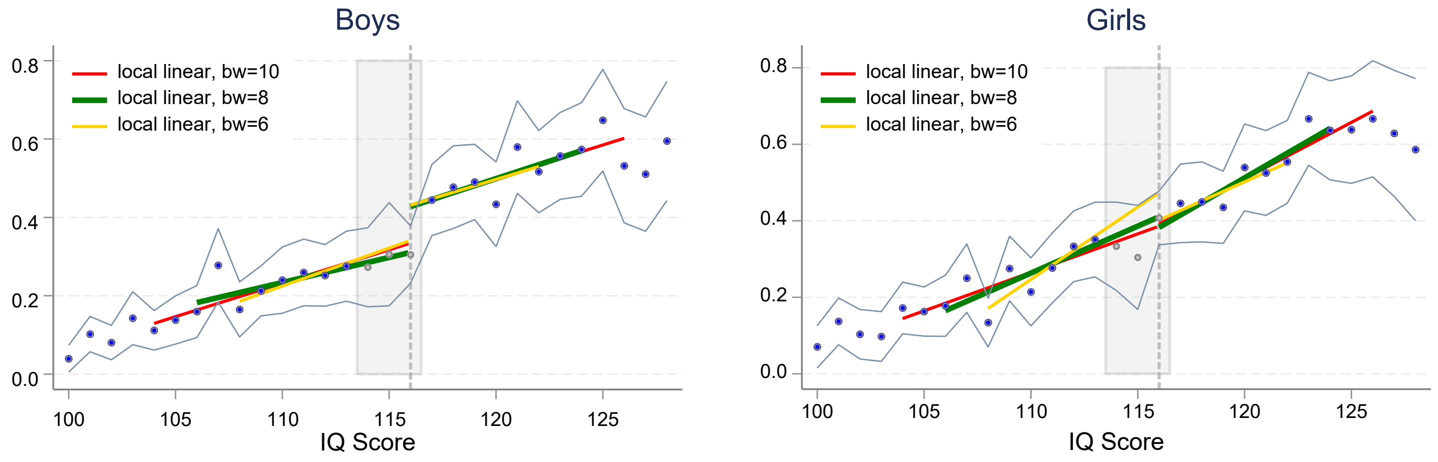
Figure 8: Participation in Gifted/High Achiever Classrooms in Grade 4, Gifted and Non-Gifted Compliers



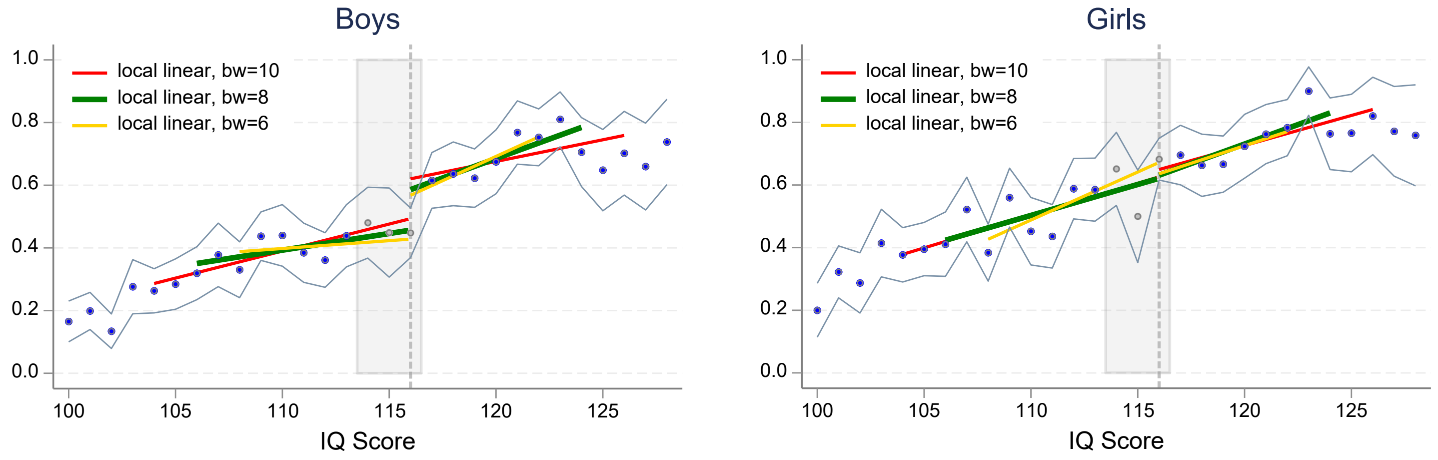
Notes: This figure shows the estimated probabilities of being in a Gifted/High Achiever Classroom in 4th grade for gifted and non-gifted compliers. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.)

Figure 9: Reduced-Form Relationships, Advanced Course Selection

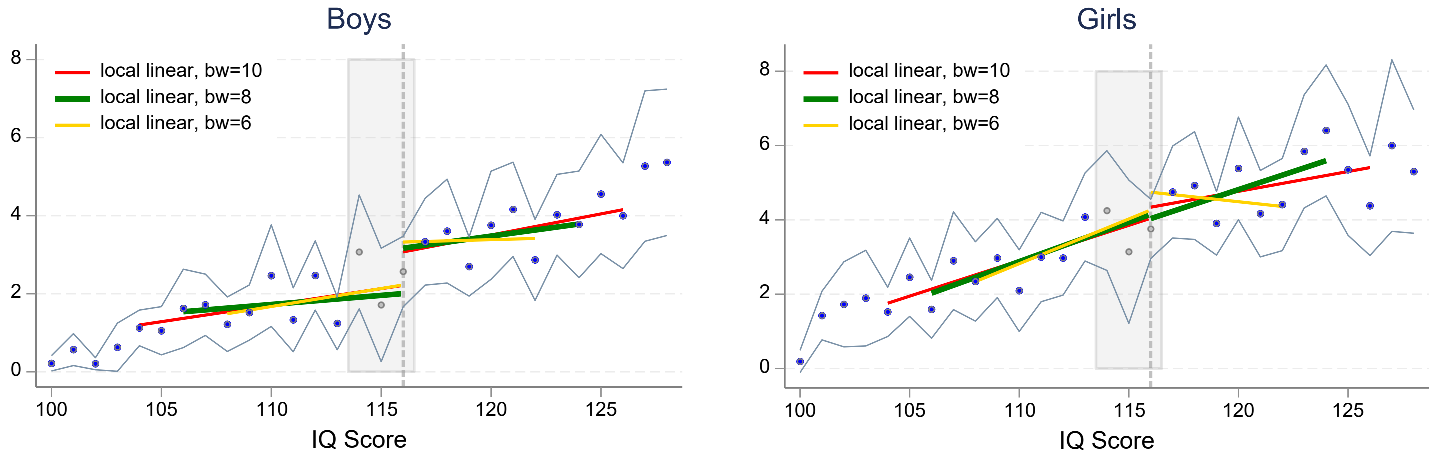
A. Took Accelerated Math Track (“GEM”) in 6th Grade



B. Took Algebra 1 Before High School



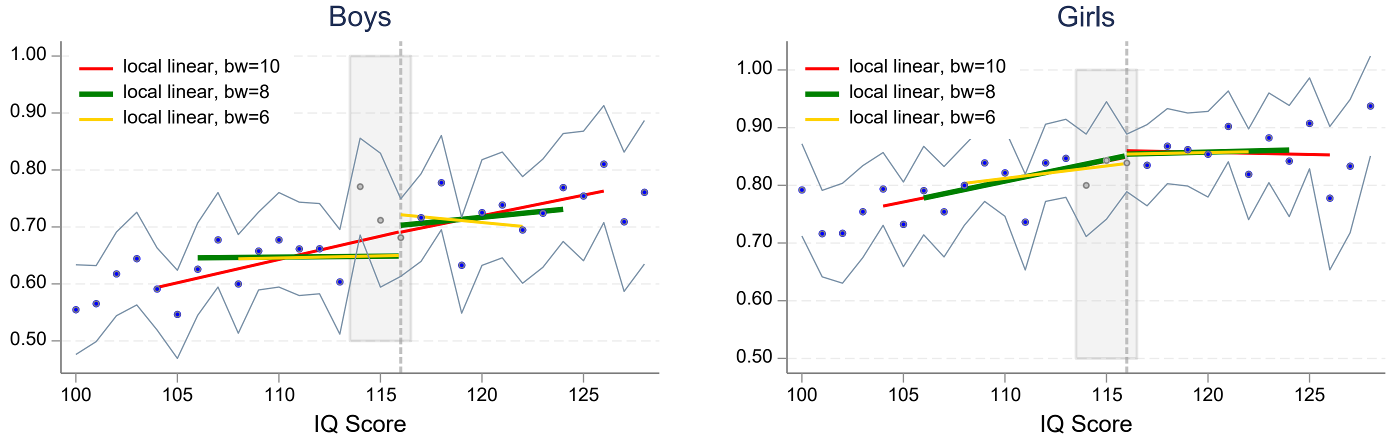
C. Number of AP Courses Taken



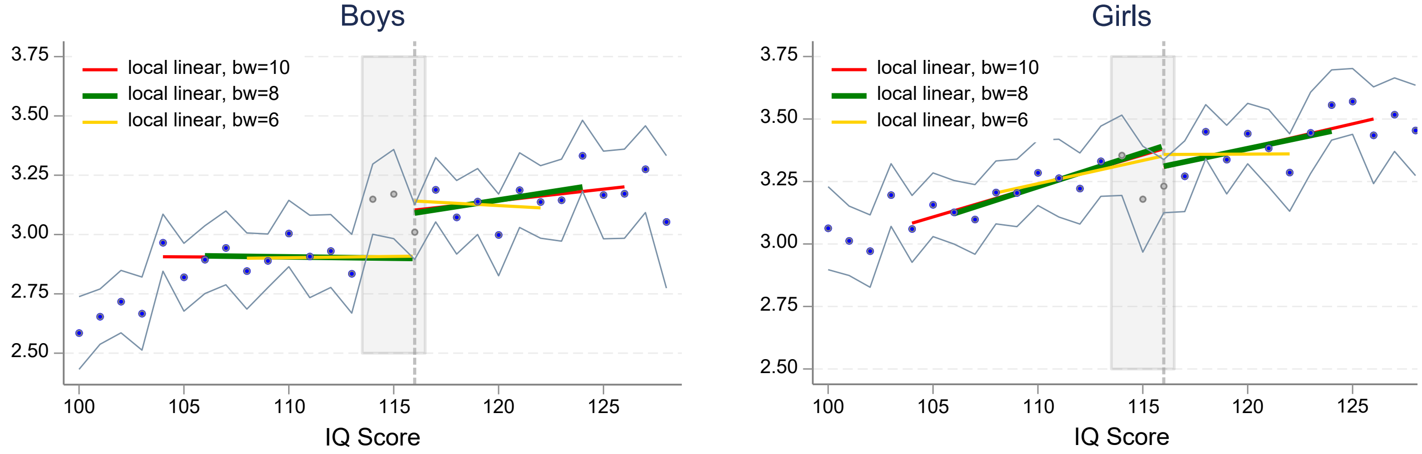
Notes: This figure plots sample means (blue dots) of the dependent variable at each value of IQ, for boys (left) and girls (right), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the “donut” region (114-116); the means for IQ scores in this region are indicated by open circles. The titles in each row indicate the dependent variable.

Figure 10: Reduced-Form Relationships, Middle-School Non-Cognitive Index and High-School Grades

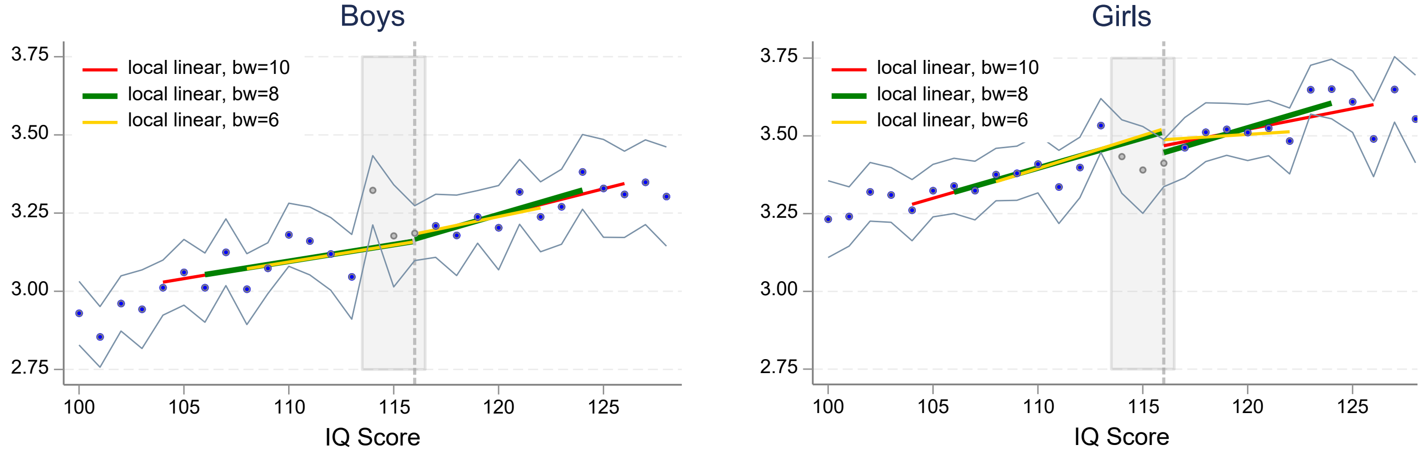
A. No Suspensions in Middle School



B. High School GPA – Math Courses

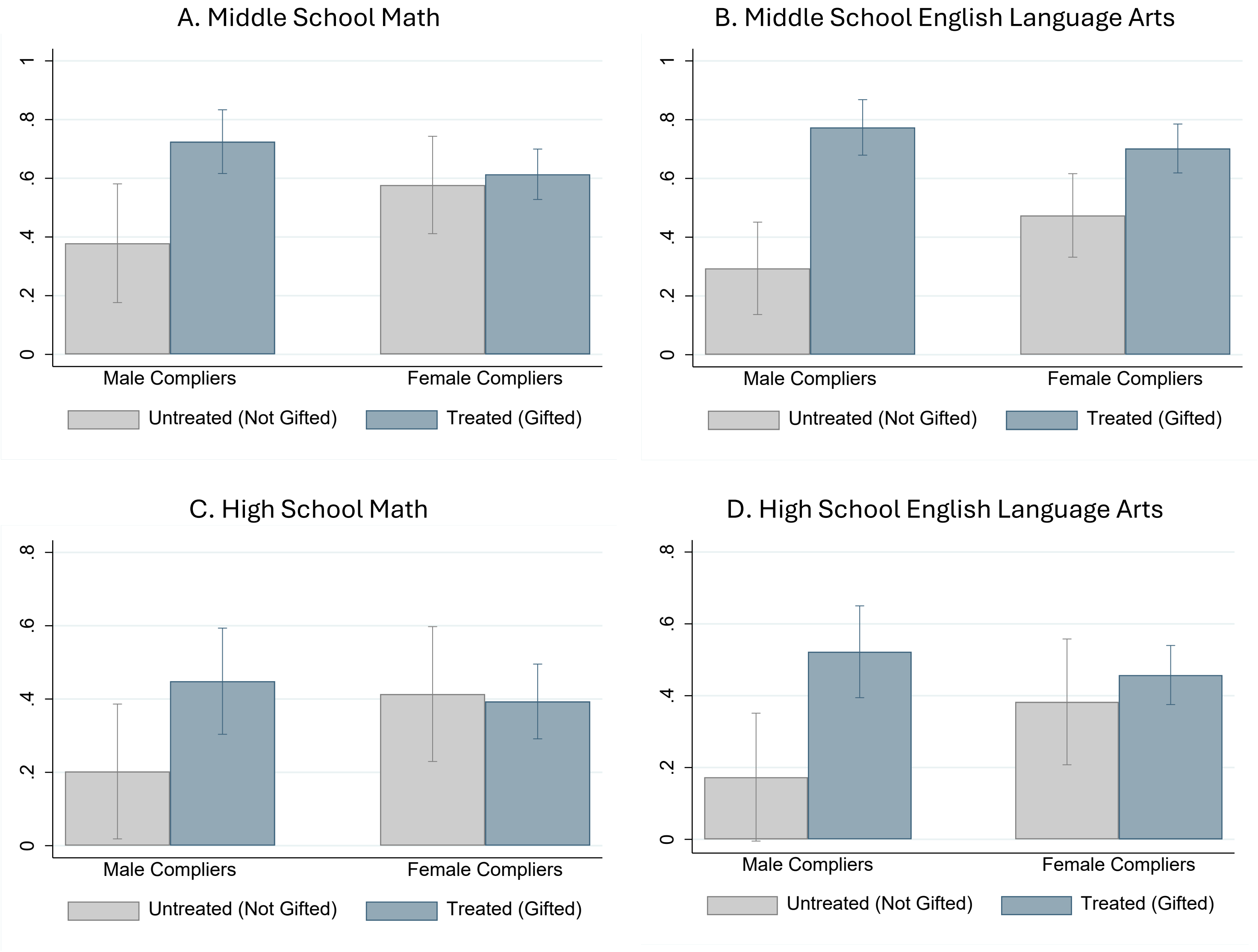


C. High School GPA – Overall



Notes: This figure plots sample means (blue dots) of the dependent variable at each value of IQ, for boys (left) and girls (right), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. The titles in each row indicate the dependent variable.

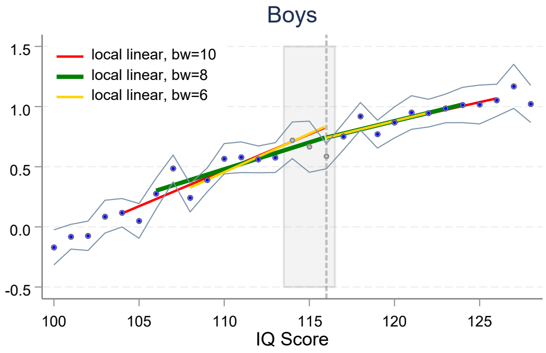
Figure 11: Classroom Peer Quality in Middle and High School, Gifted and Non-Gifted Compliers



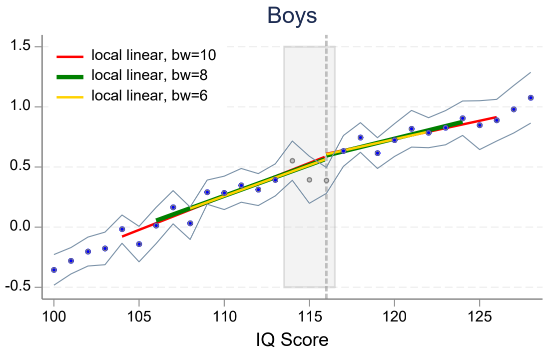
Notes: This figure shows the estimated average quality (as measured by average test scores from grades 3-5) of classroom peers for gifted and non-gifted compliers. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.)

Figure 12: Reduced-Form Relationships, Standardized Test Scores

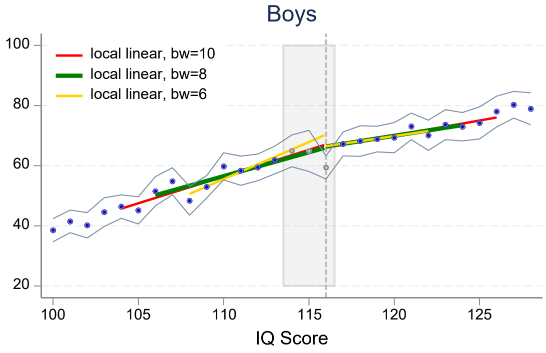
A. Grades 6-8 Test Scores (Math)



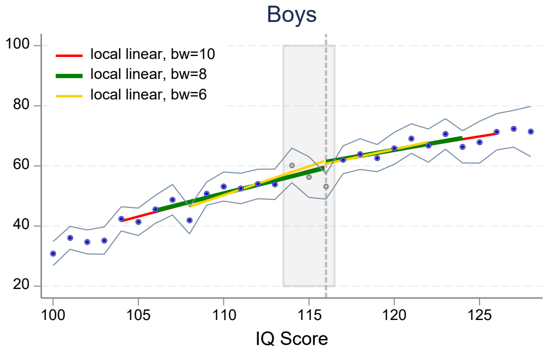
B. Grades 6-8 Test Scores (Reading)



C. PSAT Score (Math)



D. PSAT Score (Verbal)



Notes: This figure plots sample means (blue dots) of the dependent variable at each value of IQ, for boys (left) and girls (right), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. The titles in each row indicate the dependent variable.

Table 1. Summary Statistics, Students in the District

	(1)	(2)	(3)	(4)	(5)	(6)
		Plan B + IQ Scores 106-124			Compliers	
	All Students	Pooled	Boys	Girls	Boys	Girls
Student Characteristics						
White	0.32	0.18	0.17	0.18	0.15	0.28
					(0.04)	(0.05)
Black	0.35	0.37	0.34	0.39	0.32	0.36
					(0.06)	(0.06)
Hispanic	0.26	0.37	0.40	0.34	0.38	0.29
					(0.06)	(0.05)
FRL	0.49	0.80	0.78	0.83	0.78	0.84
					(0.05)	(0.05)
ELL	0.08	0.05	0.06	0.04	0.03	0.03
					(0.02)	(0.01)
G2 Avg. Test *z*-score	0.03	0.53	0.45	0.62	0.94	0.90
					(0.08)	(0.07)
G2 Math Test *z*-score	0.03	0.57	0.58	0.56	1.07	0.80
					(0.09)	(0.09)
G2 Reading Test *z*-score	0.04	0.49	0.32	0.68	0.81	1.00
					(0.10)	(0.08)
G3 Avg. Test *z*-score	0.06	0.59	0.54	0.65	0.86	0.91
					(0.07)	(0.06)
G3 Math Test *z*-score	0.06	0.66	0.65	0.67	0.99	0.90
					(0.08)	(0.07)
G3 Reading Test *z*-score	0.07	0.52	0.42	0.64	0.74	0.93
					(0.08)	(0.08)
G3, High Non-Cognitive Skills	0.60	0.63	0.59	0.67	0.65	0.79
					(0.06)	(0.07)
Student's School Characteristics						
Share Black	0.35	0.38	0.37	0.40	0.39	0.46
					(0.05)	(0.04)
Share Hispanic	0.26	0.26	0.27	0.26	0.24	0.24
					(0.02)	(0.02)
Share FRL	0.49	0.56	0.55	0.57	0.54	0.62
					(0.04)	(0.03)
Avg. Test z-score	0.01	-0.05	-0.04	-0.06	-0.02	-0.12
					(0.05)	(0.04)
Observations	227,754	3,526	1,879	1,647	--	--

Notes: This table presents means for fifth grade students in the District. Sample in column 1 is all students who were in 5th grade in 2003-2012. Sample in columns 2-4 is students in the Plan B sample with first IQ scores between 106-124 points (our preferred bandwidth). Sample in columns 5-6 is Plan B compliers. Plan B designation is based on FRL and ELL status at the time of an IQ exam. Since the Plan B sample of fifth graders includes students whose first IQ exam was before fifth grade, the FRL and ELL rates do not sum to unity. Complier characteristics are calculated using a linear RD specification where the outcome is specified as the interaction between the characteristic and gifted status and the sample includes students with IQ scores 104-126 (Abadie, 2002). The sample sizes in the bottom row indicate the maximal number of individuals included in the column. Information on 2nd grade test scores is available for approximately 70 percent of students.

Table 2. Regression Discontinuity Effects for On-time Graduation and College Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline (Pre-gifted) Measures			First Stage	Reduced Form	2SLS
	Grade 2, average *z*-score (Math & Reading)	Grade 3, average *z*-score (Math & Reading)	Predicted On-time Enrollment	Probability of being Identified as Gifted	Grad HS On time + Enrolled Within 1 Year	Grad HS On time + Enrolled Within 1 Year
Panel A. Boys Only						
BW 8	0.11	0.05	0.00	0.52**	0.14**	0.28**
	(0.11)	(0.09)	(0.02)	(0.05)	(0.05)	(0.10)
Obs.	1,285	1,819	1,879	1,879	1,879	1,879
BW 6	-0.04	-0.02	0.00	0.47**	0.19**	0.41**
	(0.14)	(0.12)	(0.03)	(0.05)	(0.07)	(0.16)
Obs.	1,000	1,406	1,451	1,451	1,451	1,451
BW 10	-0.01	-0.00	-0.00	0.51**	0.09*	0.18*
	(0.09)	(0.07)	(0.02)	(0.04)	(0.04)	(0.09)
Obs.	1,595	2,264	2,336	2,336	2,336	2,336
Panel B. Girls Only						
BW 8	-0.05	-0.07	-0.01	0.60**	0.03	0.05
	(0.10)	(0.08)	(0.01)	(0.04)	(0.06)	(0.09)
Obs.	1,166	1,612	1,647	1,647	1,647	1,647
BW 6	-0.14	-0.21+	-0.01	0.62**	0.04	0.07
	(0.13)	(0.11)	(0.02)	(0.05)	(0.08)	(0.12)
Obs.	915	1,264	1,294	1,294	1,294	1,294
BW 10	-0.03	-0.10	-0.03*	0.61**	0.02	0.03
	(0.10)	(0.07)	(0.01)	(0.04)	(0.05)	(0.08)
Obs.	1,440	2,008	2,048	2,048	2,048	2,048

Notes: This table presents estimates from local linear regression discontinuity specifications. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). The column label at the top indicates the dependent variable in each specification. In columns 1-5, the entries are coefficients on an indicator for having an IQ score ≥ 116. Column 6 shows 2SLS estimates for the effect of being identified as gifted using an indicator for having an IQ score ≥ 116 as an instrument for gifted status. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table 3. Regression Discontinuity Effects for Advanced Course Selection in Elementary, Middle, and High-School

	(1) In GHA Classroom	(2) GEM Eligible	(3) GEM in 6th Grade	(4) Algebra by 8th Grade	(5) # AP Courses
Panel A. Boys Only					
BW 8	0.72**	0.11	0.26*	0.30**	2.47**
	(0.10)	(0.12)	(0.10)	(0.10)	(0.88)
Obs.	1,506	1,590	1,879	1,879	902
BW 6	0.82**	0.12	0.20	0.31*	2.51+
	(0.14)	(0.17)	(0.14)	(0.15)	(1.28)
Obs.	1,176	1,236	1,451	1,451	683
BW 10	0.70**	0.08	0.22**	0.27**	1.81*
	(0.09)	(0.10)	(0.08)	(0.08)	(0.75)
Obs.	1,874	1,963	2,336	2,336	1,122
Panel B. Girls Only					
BW 8	0.49**	-0.05	-0.07	-0.04	0.26
	(0.11)	(0.11)	(0.10)	(0.10)	(0.88)
Obs.	1,360	1,413	1,647	1,647	715
BW 6	0.45**	-0.13	-0.14	-0.08	0.53
	(0.13)	(0.12)	(0.11)	(0.12)	(1.00)
Obs.	1,077	1,117	1,294	1,294	543
BW 10	0.55**	0.04	-0.04	-0.02	0.61
	(0.08)	(0.10)	(0.09)	(0.09)	(0.75)
Obs.	1,693	1,745	2,048	2,048	897

Notes: This table presents estimates from local linear regression discontinuity specifications. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). All columns show 2SLS estimates for the effect of being identified as gifted using an indicator for having an IQ score ≥ 116 as an instrument. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). The column label at the top indicates the dependent variable in each specification. Enrollment in GHA classrooms is only available for the cohorts from 2006-2012. GEM eligibility is only measured for students who have fifth grade standardized math achievement scores. The number of AP courses in high school grades is only available for the cohorts 2003-2007 and 2012. Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table 4. Regression Discontinuity Effects for Course Grades and Disciplinary Actions

	(1) G6-G8 Math GPA	(2) G6-G8 No Suspensions	(3) HS Math GPA	(4) HS Lang. Arts GPA	(5) HS Overall GPA
Panel A. Boys Only					
BW 8	0.17	0.10	0.37+	0.11	0.01
	(0.15)	(0.10)	(0.20)	(0.19)	(0.16)
Obs.	1,879	1,879	1,879	1,879	1,879
BW 6	0.10	0.15	0.50+	0.31	0.05
	(0.21)	(0.14)	(0.28)	(0.26)	(0.23)
Obs.	1,451	1,451	1,451	1,451	1,451
BW 10	0.04	-0.00	0.40*	0.05	0.02
	(0.14)	(0.09)	(0.17)	(0.15)	(0.14)
Obs.	2,336	2,336	2,336	2,336	2,336
Panel B. Girls Only					
BW 8	-0.10	0.00	-0.13	0.01	-0.11
	(0.12)	(0.08)	(0.16)	(0.11)	(0.10)
Obs.	1,647	1,647	1,647	1,647	1,647
BW 6	-0.07	0.03	0.01	0.03	-0.06
	(0.13)	(0.08)	(0.18)	(0.15)	(0.13)
Obs.	1,294	1,294	1,294	1,294	1,294
BW 10	-0.01	0.02	-0.12	0.01	-0.08
	(0.10)	(0.07)	(0.13)	(0.09)	(0.09)
Obs.	2,048	2,048	2,048	2,048	2,048

Notes: This table presents estimates from local linear regression discontinuity specifications. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). All columns show 2SLS estimates for the effect of being identified as gifted using an indicator for having an IQ score ≥ 116 as an instrument. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). The column label at the top indicates the dependent variable in each specification. Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table 5. Regression Discontinuity Effects for Middle and High School Peer Measures

	(1) G6 Chall. Courses	(2) G6-G8 Peers, All	(3) G6-G8 Peers, Math Courses	(4) G6-G8 Peers, Reading Courses	(5) HS Peers, All	(6) HS Peers, Math Courses	(7) HS Peers, Reading Courses
Panel A. Boys Only							
BW 8	0.89+	0.32**	0.35**	0.48**	0.20*	0.25*	0.35**
	(0.49)	(0.08)	(0.11)	(0.09)	(0.10)	(0.12)	(0.11)
Obs.	729	1,879	1,879	1,839	902	902	901
BW 6	1.26+	0.31*	0.23	0.52**	0.08	0.10	0.25
	(0.71)	(0.13)	(0.17)	(0.14)	(0.14)	(0.17)	(0.15)
Obs.	574	1,451	1,451	1,422	683	683	682
BW 10	0.71+	0.24**	0.22*	0.40**	0.12+	0.11	0.24**
	(0.42)	(0.07)	(0.09)	(0.08)	(0.07)	(0.09)	(0.09)
Obs.	897	2,336	2,336	2,283	1,122	1,122	1,121
Panel B. Girls Only							
BW 8	1.60**	0.14*	0.04	0.23**	0.07	-0.02	0.07
	(0.50)	(0.07)	(0.09)	(0.08)	(0.08)	(0.10)	(0.09)
Obs.	678	1,647	1,646	1,624	715	715	715
BW 6	1.49*	0.04	-0.09	0.11	0.08	-0.04	0.09
	(0.61)	(0.08)	(0.12)	(0.10)	(0.09)	(0.11)	(0.10)
Obs.	553	1,294	1,293	1,279	543	543	543
BW 10	1.25**	0.12*	0.02	0.22**	0.02	-0.05	0.04
	(0.38)	(0.06)	(0.08)	(0.07)	(0.06)	(0.08)	(0.07)
Obs.	840	2,048	2,047	2,015	897	897	897

Notes: This table presents estimates from linear regression discontinuity specifications. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). All columns show 2SLS estimates for the effect of being identified as gifted using an indicator for having an IQ score ≥ 116 as an instrument. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). The column label at the top indicates the dependent variable in each specification. The grade 6 challenging course survey question in Column 1 is only measured for students in the 2008-2012 cohorts that participated in the survey. The peer measures for high school grades (Columns 5-7) are only available for the cohorts 2003-2007 and 2012. Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table 6. Regression Discontinuity Effects for Middle and High-School Test Scores

	(1) G6-G8 Math Scores	(2) G6-G8 Reading Scores	(3) PSAT Math (Percentile)	(4) PSAT Reading (Percentile)
Panel A. Boys Only				
BW 8	-0.01	0.05	0.60	3.63
	(0.15)	(0.17)	(5.86)	(6.14)
Obs.	1,878	1,879	1,709	1,709
BW 6	-0.21	0.08	-7.74	-2.12
	(0.23)	(0.25)	(8.28)	(8.43)
Obs.	1,450	1,451	1,329	1,329
BW 10	-0.16	0.04	-2.00	3.23
	(0.14)	(0.14)	(5.56)	(5.20)
Obs.	2,335	2,336	2,117	2,117
Panel B. Girls Only				
BW 8	-0.09	-0.21	-4.06	-2.79
	(0.13)	(0.14)	(4.45)	(4.83)
Obs.	1,647	1,647	1,519	1,519
BW 6	-0.22	-0.27	-9.25+	-5.09
	(0.16)	(0.17)	(5.25)	(6.10)
Obs.	1,294	1,294	1,187	1,187
BW 10	-0.09	-0.03	-2.15	1.80
	(0.09)	(0.08)	(3.20)	(3.04)
Obs.	4,383	4,384	3,989	3,989

Notes: This table presents estimates from linear regression discontinuity specifications. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). All columns show 2SLS estimates for the effect of being identified as gifted using an indicator for having an IQ score ≥ 116 as an instrument. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). The column label at the top indicates the dependent variable in each specification. Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Table 7. Gender Gaps for Non-gifted Compliers and the Effects of Gifted Status

	Non-Gifted Complier Outcomes		Effect of Gifted Status:	
	Female Y(0) Mean	Female - Male Y(0) Gap	Effect on Boys	Differential Effect on Boys
	(1)	(2)	(3)	(4)
A. Non-Cognitive Outcomes				
Gifted/High-Achieving Classroom Enrollment	0.43	0.17	0.72**	0.23
GEM Enrollment, 6th Grade	0.49	0.23	0.26*	0.34**
G6-G8 Peers, Math Courses (σ units)	0.58	0.20	0.35**	0.31**
G6-G8 Peers, Reading Courses (σ units)	0.47	0.18	0.48**	0.25**
Algebra Enrollment, by 8th Grade	0.67	0.25	0.30**	0.33**
G6-G8, No Suspensions	0.85	0.18	0.10	0.10
HS Peers, Math Courses (σ units)	0.41	0.21	0.25*	0.27*
HS Peers, Reading Courses (σ units)	0.38	0.21	0.35**	0.27*
HS Math GPA	3.46	0.65	0.37+	0.50**
HS Lang. Arts GPA	3.68	0.46	0.11	0.11
HS GPA	3.58	0.34	0.01	0.12
Number of AP Courses	3.94	1.81	2.47**	2.21*
B. Cognitive Outcomes				
G6-G8 Test Scores (σ units)	1.04	0.25	0.04	0.20
G6-G8 Math Test Scores (σ units)	0.96	0.04	-0.01	0.08
G6-G8 Reading Test Scores (σ units)	1.12	0.43	0.05	0.27
PSAT (Percentile)	71.69	6.32	0.45	2.54
PSAT Math (Percentile)	71.32	0.01	0.60	4.66
PSAT Reading (Percentile)	75.93	13.9	3.63	6.42
C. College				
Grad HS On Time + Enroll Within 1 Year	0.74	0.28	0.28**	0.23

Notes: Each row reports summary statistics for the education outcome described in the first column for students in our main analysis sample. Column 1 reports estimated means of the outcome for girl compliers if they had not been exposed to treatment (i.e., the estimated means of Y(0), in standard notation). Column 2 reports the difference in estimated means of the complier untreated outcomes (i.e., the gender gap in Y(0) between girls and boys). Column 3 shows the estimated effect for boys of achieving gifted status by the end of 5th grade on the outcome. Column 4 shows the difference in the estimated effects for boys relative to girls. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). Statistical significance of estimated effects in columns 3-4 is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

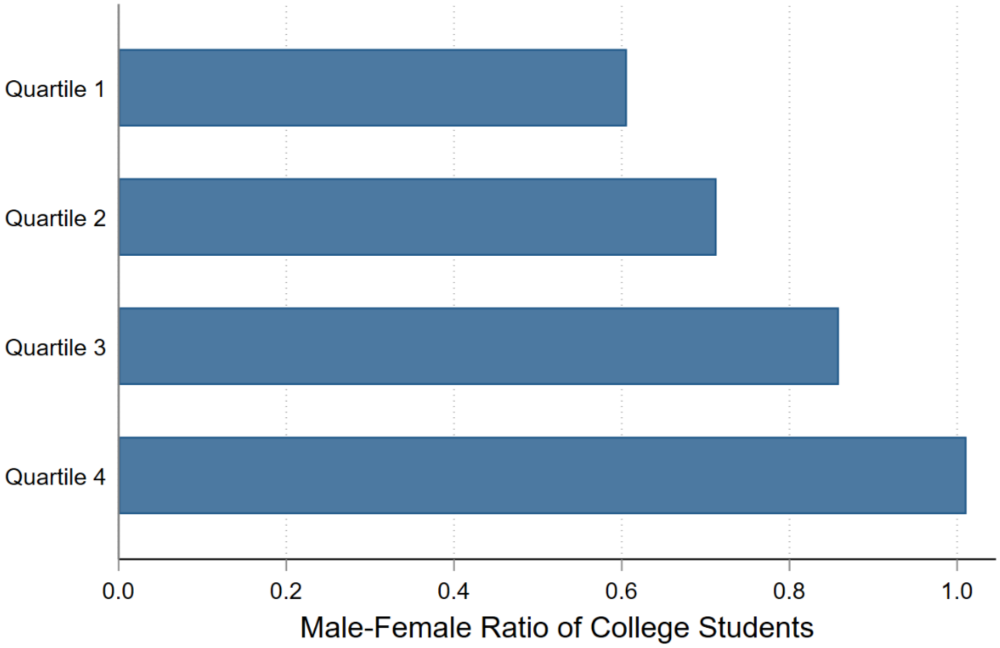
Online Appendix

Can Gifted Education Help Higher-Ability Students from Disadvantaged Backgrounds?

David Card, Eric Chyn, and Laura Giuliano

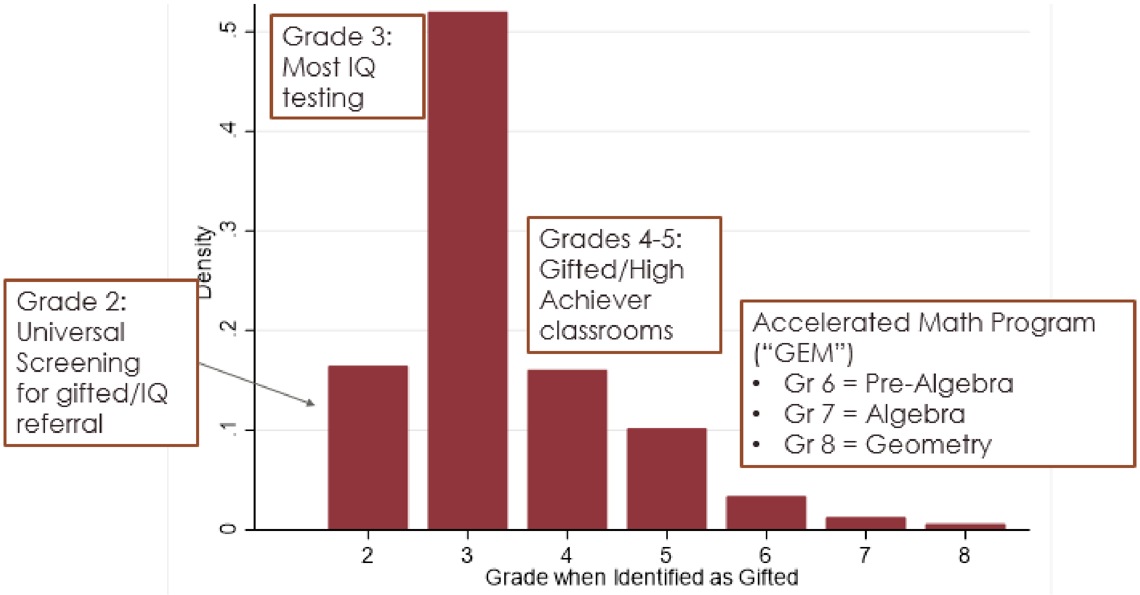
Appendix A: Additional Figures and Tables

Figure A1: Male-Female Ratio of College Freshmen by Family Income Quartile



Source: U.S. Department of Education, National Center for Education Statistics, 1995-96 National Postsecondary Student Aid Study (NPSAS:96), NPSAS:2000, NPSAS:04, NPSAS:08, NPSAS:12 and NPSAS:16. Computation by NCES TrendStats on 8/28/2021.

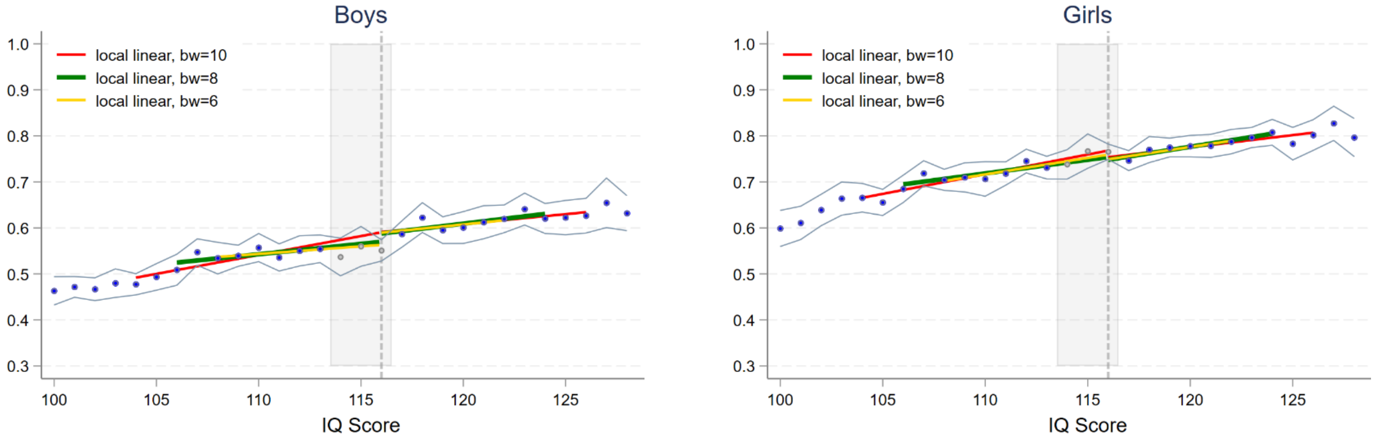
Figure A2: Timing of Gifted Screening, Identification, and Relevant Programs



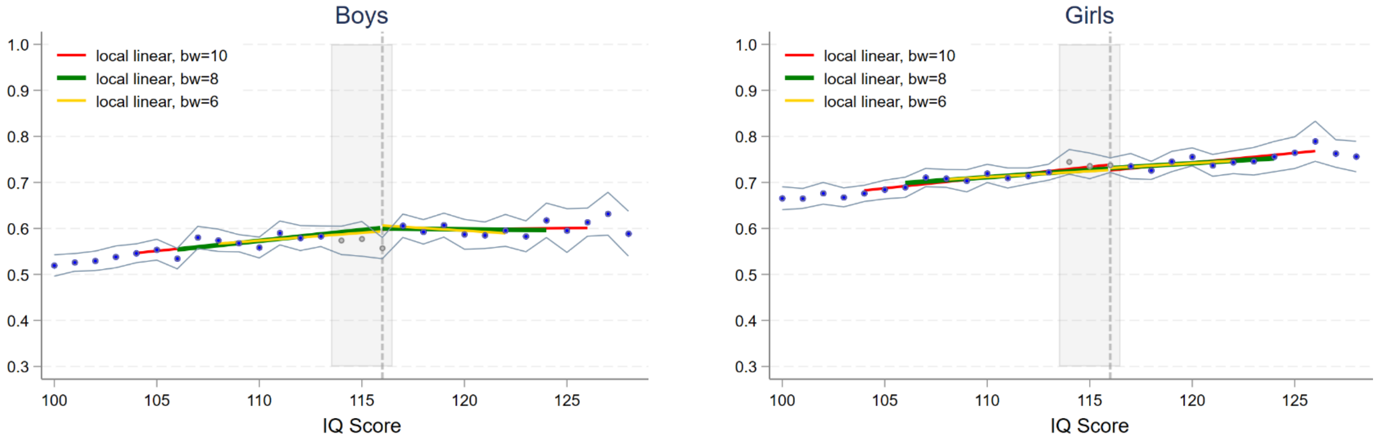
Notes: This figure is a histogram of the grade when a student was identified as gifted for the sample of all gifted students who were in 5th grade in the District between 2003-2012 (the cohorts in our analysis sample) and who were FRL participants or English Language Learners (i.e., "Plan B" eligible) at the time of their first IQ test. Information on the timing of relevant educational policies and programs is also illustrated.

Figure A3: Reduced-Form Relationships, Predicted On-Time Graduation and College Enrollment Using Second-Grade Test Scores

A. Using 2nd-grade SAT scores



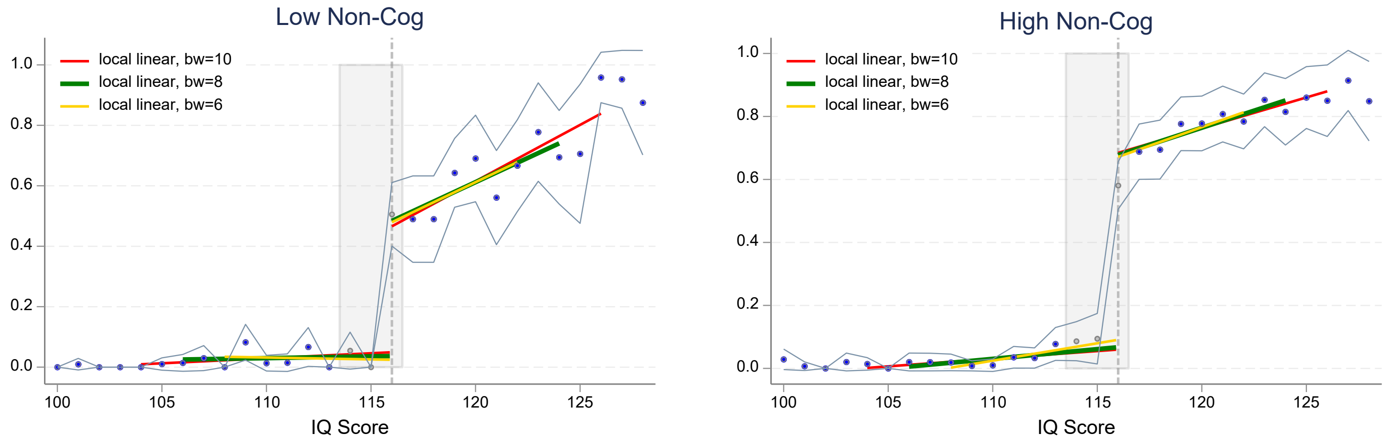
B. Using 2nd-grade Nonverbal Ability Index



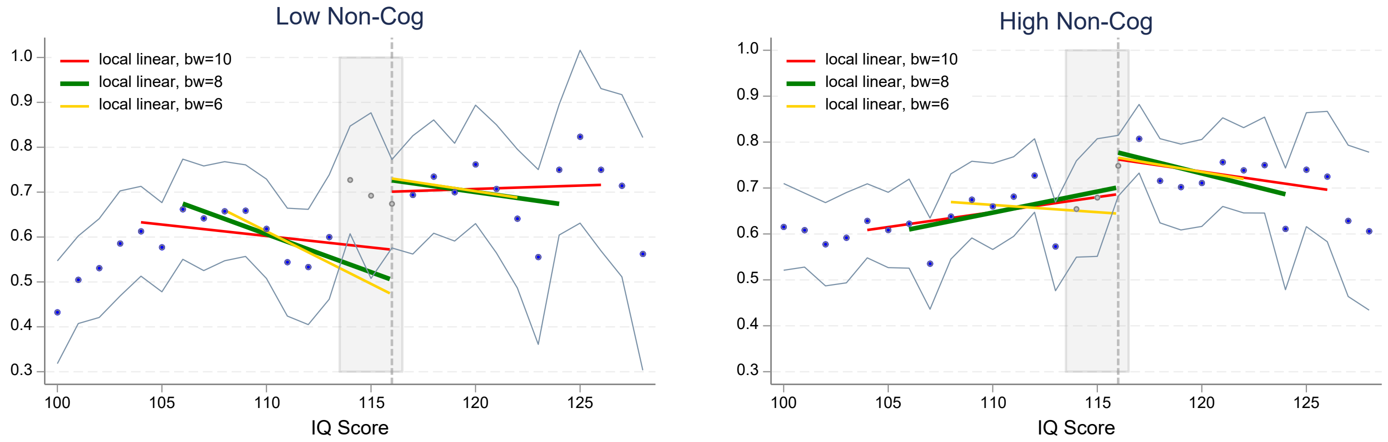
Notes: This figure plots sample means (blue dots) of predicted enrollment at each value of IQ, for boys (Panel A) and girls (Panel B), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. The dependent variable is predicted probability of graduating on time and enrolling in college from models fit separately to boys and girls with IQ scores in the range from 100 to 115. The prediction models are similar to those used in Panel B of Figure 4 but replace 3rd-grade scores with either 2nd-grade SAT scores in math and reading (Panel A) or 2nd-grade Nonverbal Ability Index based on the test used for gifted screening in 2nd grade (Panel B). The models also include students' responses in grade 3 to a survey question about how much they enjoy learning in their school; the number of internal and external suspensions in grade 3; indicators for student race/ethnicity, FRL status, ELL status, and cohort; average test scores, fraction FRL and fraction nonwhite of the school where the student is enrolling in grade 5; and the median household income of the student's neighborhood.

Figure A4: First-Stage and Reduced-Form Relationships for On-Time College Enrollment, Boys and Girls with High vs. Low Non-Cognitive Skills in Grade 3

A. First Stage Relationships for Gifted Status

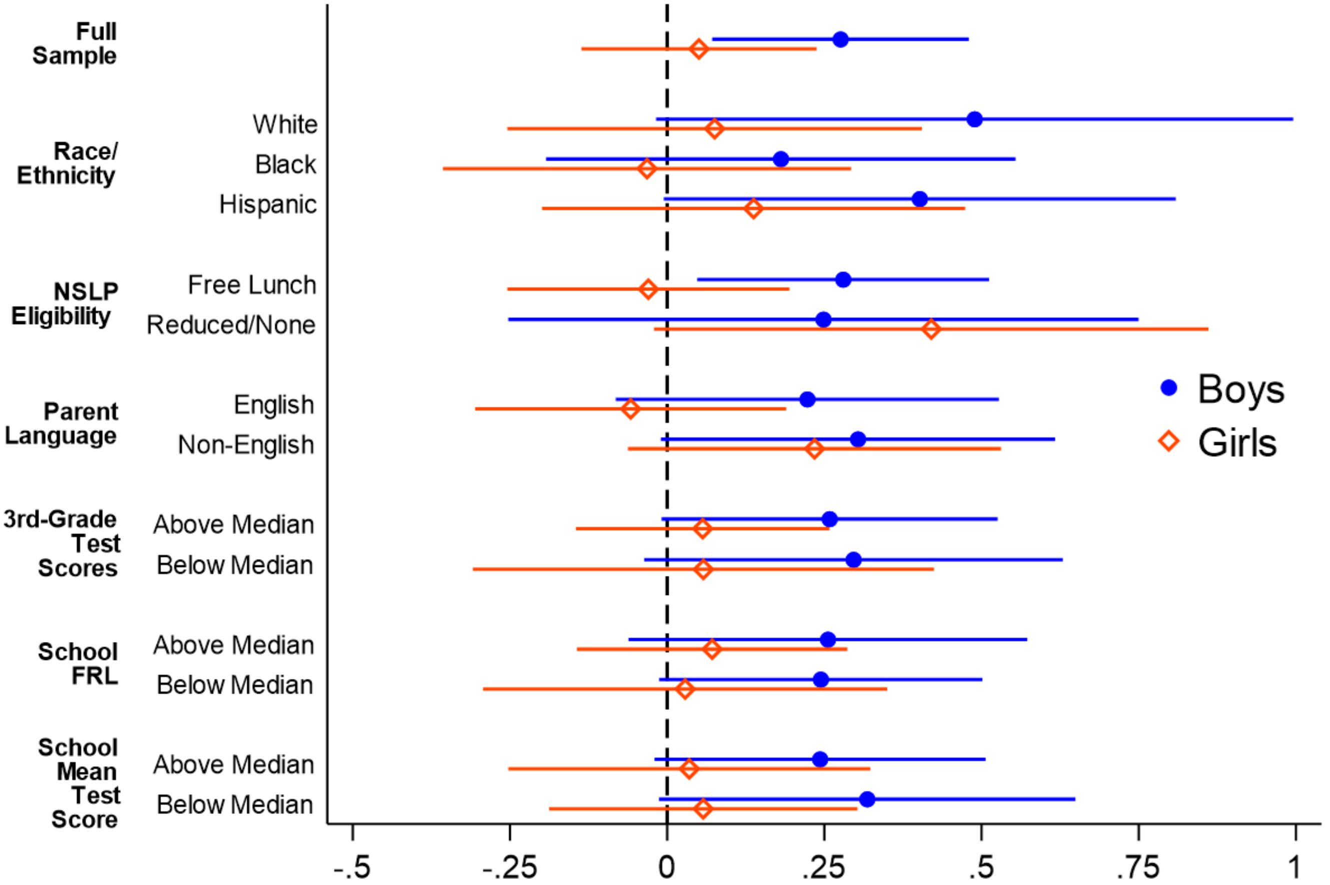


B. Reduced-Form Relationships for On-Time College Enrollment



Notes: This figure plots sample means (blue dots) of the dependent variable at each value of IQ for two samples along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. In Panel A, the dependent variable is an indicator for gifted status (as of the end of 5th grade); in Panel B, it is an indicator for graduating high school on time and enrolling in college within one year. Each panel shows plots for two samples: the figures on the left use boys and girls with low non-cognitive skills in 3rd grade and those on the right use boys and girls with high non-cognitive skills. As in Figure 7, students are classified as having high non-cognitive skills if they had no disciplinary actions in 3rd grade and if, in response to a District-administered survey, they "strongly agreed" that they enjoyed learning at their school.

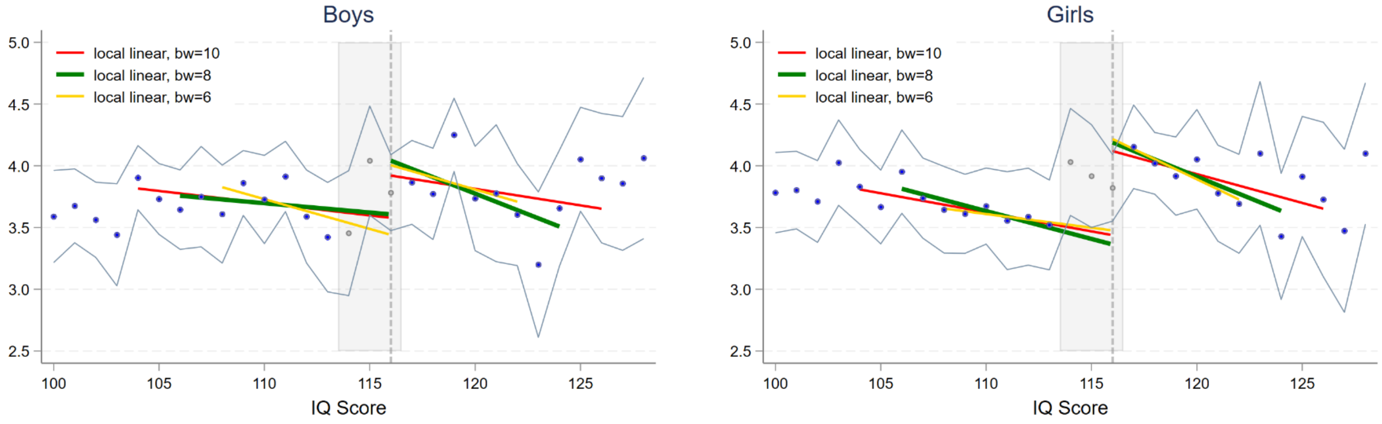
Figure A5: Other Potential Sources of Treatment Effect Heterogeneity



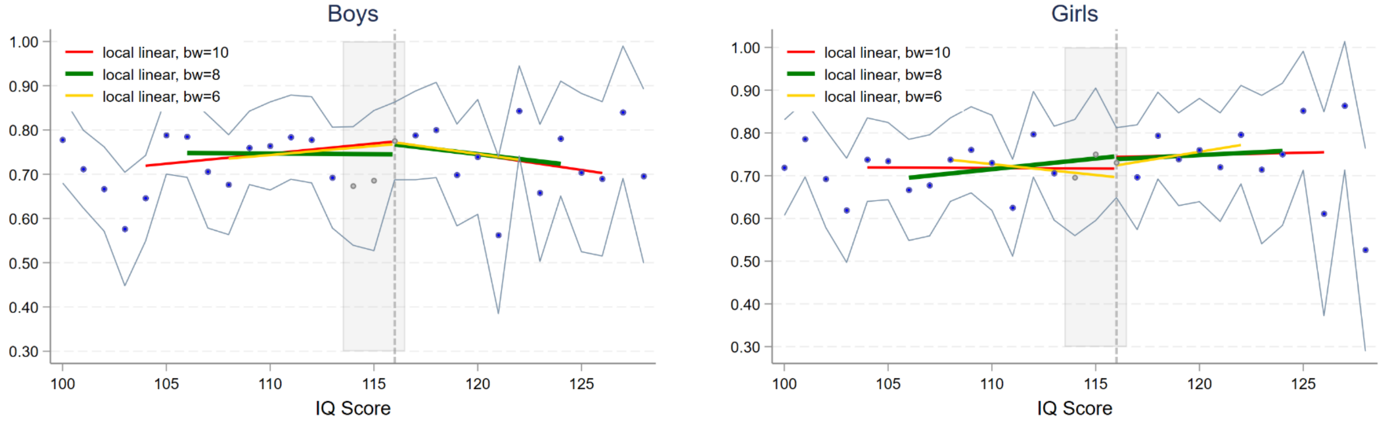
Notes: This figure reports 2SLS estimates from local linear fuzzy RD models where the endogenous variable is gifted status. All models use a symmetric bandwidth of 8 (IQ scores between 106-124) and exclude IQ scores that fall in the “donut” region (scores between 114-116). Point estimates for the effect of gifted status are presented by subgroup (as indicated in the row labels) separately for boys (blue, circle marker) and girls (orange, diamond marker). The lines represent 95% confidence intervals for the respective estimates.

Figure A6: Reduced-Form Relationships, Survey Response on Guidance with Selecting Challenging Courses

A. Response on Scale from 1 (Strongly Disagree) to 5 (Strongly Agree)



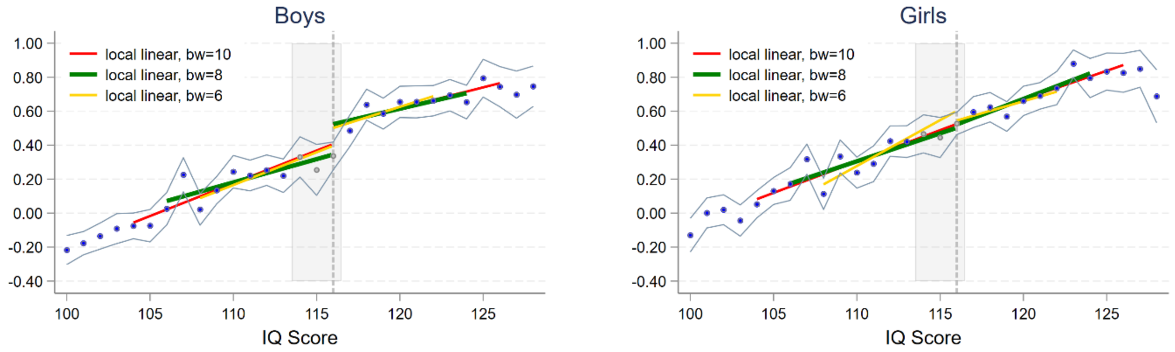
B. Probability of Non-Missing Response



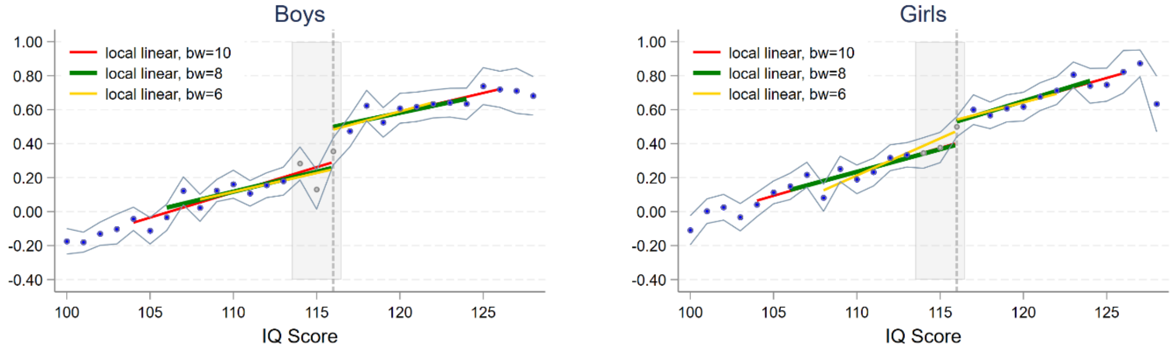
Notes: This figure plots sample means (blue dots) of the dependent variable at each value of IQ, for boys (left) and girls (right), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. The dependent variable in the top row (Panel A) is the survey response of 6th graders, collected on a 5-point Likert scale measuring agreement with the statement: "This year, school staff helped me to select high level courses that challenge my abilities." The student questionnaire was only collected in the 2008-2011 5th grade cohorts, so this analysis is based only on students in these four cohorts who had non-missing responses. In Panel B, the dependent variable is an indicator for having a non-missing response; the analysis sample is again limited to the four cohorts who received the survey.

Figure A7: Reduced-Form Relationships, Peer Quality in Middle and High School Courses

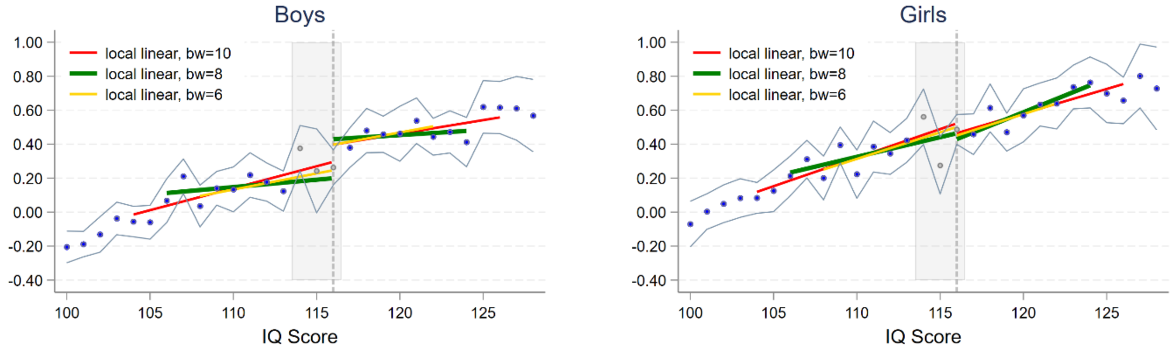
A. Middle School Math Courses



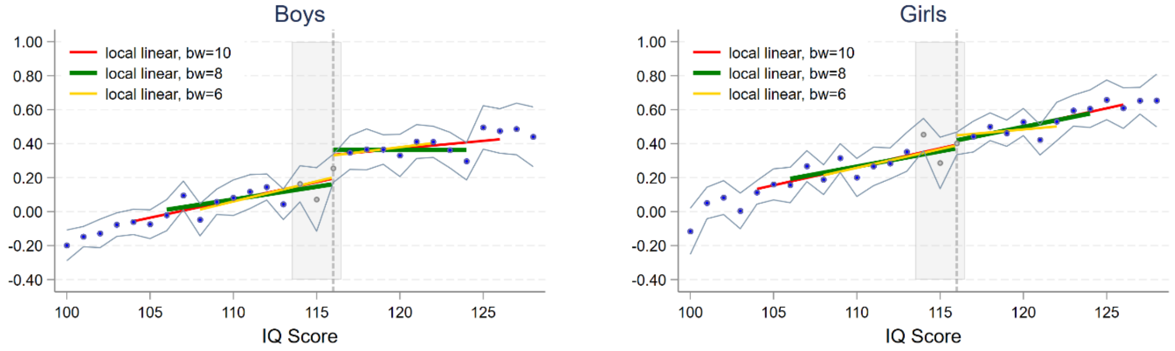
B. Middle School English Language Arts Courses



C. High School Math Courses



D. High School English Language Arts Courses



Notes: This figure plots sample means (blue dots) of classmate (peer) quality, as represented by their mean test scores in grades 3-5, in the set of classed indicated in the panel heading, at each value of IQ, for boys (left) and girls (right), along with 95% confidence intervals for the means (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 10 (red), 8 (green), or 6 (yellow). All models are estimated excluding IQ scores that fall in the "donut" region (114-116); the means for IQ scores in this region are indicated by open circles. The peer measures for high school grades in panels C and D are only available for the cohorts 2003-2007 and 2012.

Appendix Table A1. Characteristics of Compliers with Participation in Gifted/High Achiever Classroom in Grade 4

	(1) Plan B Gifted; (Not High Achievers)	(2) Non-Gifted High Achievers	(3) Minority Non-Gifted High Achievers
Own characteristics:			
Grade 3 reading	0.62	0.90	0.84
Grade 3 math	0.81	0.84	0.64
Female	0.44	0.52	0.58
Black	0.44	0.22	0.51
Hispanic	0.29	0.21	0.49
FRL	0.91	0.29	0.04
ELL	0.19	0.01	0.03
Neighborhood Median Household Income	48.3	64.5	54.0
Non-Verbal Ability Index (NNAT)	124.9	86.6	93.1
High 3rd-Grade Non-Cog Skills	0.61	0.69	0.69
Classroom characteristics:			
Avg Gr 3 reading of Gr 4 classroom	0.65	0.95	0.87
Avg Gr 3 math of Gr 4 classroom	0.67	0.92	0.78
Classroom % gifted	0.53	0.31	0.26

Notes: This table presents estimates of the mean characteristics of gifted and non-gifted compliers with the treatment defined as being placed in a separate “gifted/high achiever” classroom in grade 4, as well as mean characteristics of their classroom peers. Column 1 shows characteristics of the compliers in our design for disadvantaged gifted students (i.e., with IQ ≥ 116) but whose 3rd-grade test scores were below the “high achiever” threshold for their school/cohort. Column 2 shows characteristics of the compliers in the RD design used in Card and Giuliano (2016b) (i.e., non-gifted “high achievers” whose 3rd-grade standardized test scores ranked relatively high within their school/cohort). Column 3 shows characteristics for the subset of compliers in Card and Giuliano (2016b) who were minorities.

Appendix Table A2. Sample Selection Tests

	(1)	(2)	(3)
		Sample	
	Pooled	Boys	Girls
BW 8	-0.01	-0.05	0.03
	(0.03)	(0.05)	(0.04)
Mean Below Threshold	0.731	0.742	0.715
Obs.	5,022	2,749	2,273
BW 6	-0.01	-0.02	0.01
	(0.05)	(0.07)	(0.06)
Mean Below Threshold	0.735	0.737	0.730
Obs.	3,889	2,110	1,779
BW 10	0.01	-0.03	0.06
	(0.03)	(0.04)	(0.04)
Mean Below Threshold	0.713	0.726	0.696
Obs.	6,273	3,448	2,825

Notes: This table presents regression discontinuity estimates of the effect of having IQ ≥116 on an indicator for being included in our estimation sample, among students who were IQ tested by 5th grade. The column label at the top indicates the sample. We show estimates for 3 ranges of IQ: (i) IQ scores in the range 106-124 (BW 8); (ii) 108-122 (BW 6); (iii) 104-126 (BW 10). All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). The means of the outcome for students below the 116-point threshold are estimated from the intercepts of the same models. Standard errors are clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Appendix Table A3. Summary Statistics on All Student Outcomes

	(1)	(2)	(3)	(4)	(5)
	Plan B + IQ Scores 106-124			Non-Gifted Compliers	
	Pooled	Boys	Girls	Boys	Girls
Panel A. Non-Cognitive Outcomes					
Gifted/High-Achieving Classroom Enrollment	0.42	0.41	0.44	0.26	0.43
GEM Enrollment, 6th Grade	0.35	0.34	0.36	0.26	0.49
G6-G8 Peers, Math Courses (σ units)	0.31	0.27	0.36	0.28	0.42
G6-G8 Peers, Reading Courses (σ units)	0.31	0.26	0.36	0.29	0.43
Algebra Enrollment, by 8th Grade	0.52	0.48	0.56	0.42	0.67
G6-G8, No Suspensions	0.75	0.68	0.83	0.67	0.85
HS Peers, Math Courses (σ units)	0.23	0.17	0.30	0.18	0.27
HS Peers, Reading Courses (σ units)	0.22	0.16	0.30	0.17	0.28
HS Math GPA	3.14	3.01	3.29	2.81	3.46
HS Lang. Arts GPA	3.41	3.24	3.60	3.22	3.68
HS GPA	3.29	3.16	3.45	3.24	3.58
Number of AP Courses	3.03	2.50	3.70	2.13	3.94
B. Cognitive Outcomes					
G6-G8 Test Scores (σ units)	0.62	0.55	0.70	0.79	1.04
G6-G8 Math Test Scores (σ units)	0.69	0.64	0.75	0.92	0.96
G6-G8 Reading Test Scores (σ units)	0.55	0.45	0.68	0.69	1.12
PSAT (Percentile)	58.98	56.23	62.08	65.37	71.69
PSAT Math (Percentile)	62.56	61.89	63.32	71.31	71.32
PSAT Reading (Percentile)	60.00	56.73	63.67	62.03	75.93
C. College					
Grad HS On Time + Enroll Within 1 Yr.	0.67	0.62	0.72	0.46	0.74
Observations	3,526	1,879	1,647	--	--

Notes: This table presents means of the main outcome variables used in the paper for the samples described in columns 2-6 of Table 1. Entries in columns 1-3 are simple means. Entries in columns 4-5 are complier means, calculated using a linear RD specification where the outcome is specified as the interaction between the characteristic and an indicator for not being gifted and the sample includes students with IQ scores 104-126 (Abadie, 2002). The sample sizes in the bottom row indicate the maximal number of individuals included in the column.

Appendix Table A4. Additional Balance Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Baseline Student Characteristics									Baseline School Characteristics			
	Age	FRL	ELL	White	Black	Hispanic	G3, Avg. Reading *z*-score	G3, Avg. Math *z*-score	G3, High Non-Cognitive Status	Share Non-White	Share FRL	Median HH Income	Avg. Test Score
Panel A. Boys Only													
BW 8	0.03	-0.01	-0.01	-0.02	-0.08	0.09	0.02	0.08	-0.00	-0.01	-0.03	-0.53	0.02
	(0.05)	(0.05)	(0.03)	(0.04)	(0.06)	(0.06)	(0.10)	(0.09)	(0.01)	(0.02)	(0.03)	(2.03)	(0.04)
Obs.	1,879	1,879	1,879	1,879	1,879	1,879	1,818	1,817	1,267	1,879	1,879	1,867	1,879
BW 6	0.05	0.01	-0.03	-0.09	-0.06	0.11	-0.03	-0.00	-0.00	0.01	-0.02	-1.04	0.00
	(0.06)	(0.06)	(0.04)	(0.06)	(0.07)	(0.07)	(0.14)	(0.13)	(0.01)	(0.03)	(0.04)	(2.39)	(0.05)
Obs.	1,451	1,451	1,451	1,451	1,451	1,451	1,405	1,405	987	1,451	1,451	1,441	1,451
BW 10	0.03	0.03	-0.01	-0.02	-0.05	0.03	0.01	-0.02	-0.00	-0.01	-0.02	-1.28	-0.01
	(0.04)	(0.04)	(0.02)	(0.03)	(0.05)	(0.05)	(0.09)	(0.08)	(0.01)	(0.02)	(0.02)	(1.47)	(0.03)
Obs.	2,336	2,336	2,336	2,336	2,336	2,336	2,263	2,260	1,580	2,336	2,336	2,323	2,336
Panel B. Girls Only													
BW 8	-0.00	0.02	-0.01	0.01	0.01	-0.00	-0.07	-0.07	0.01*	0.01	0.01	0.23	0.03
	(0.05)	(0.05)	(0.02)	(0.05)	(0.06)	(0.06)	(0.10)	(0.09)	(0.00)	(0.02)	(0.03)	(2.01)	(0.04)
Obs.	1,647	1,647	1,647	1,647	1,647	1,647	1,611	1,611	1,174	1,647	1,647	1,641	1,647
BW 6	0.00	0.01	-0.02	-0.06	0.07	0.01	-0.20	-0.22+	0.01	0.03	0.02	1.82	0.02
	(0.06)	(0.06)	(0.03)	(0.06)	(0.07)	(0.08)	(0.13)	(0.12)	(0.00)	(0.03)	(0.04)	(2.58)	(0.05)
Obs.	1,294	1,294	1,294	1,294	1,294	1,294	1,263	1,263	934	1,294	1,294	1,288	1,294
BW 10	0.05	0.00	-0.01	-0.00	0.01	0.03	-0.11	-0.10	0.00	0.02	0.03	-1.59	-0.00
	(0.04)	(0.04)	(0.02)	(0.04)	(0.05)	(0.05)	(0.09)	(0.08)	(0.00)	(0.02)	(0.03)	(1.71)	(0.04)
Obs.	2,048	2,048	2,048	2,048	2,048	2,048	2,006	2,006	1,460	2,048	2,048	2,041	2,048

Notes: This table presents estimates from local linear regression discontinuity specifications of the effect of having an IQ score ≥ 116 on the outcome indicated by the column heading. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Appendix Table A5. Discontinuities in Probability of Graduating On-Time and Entering College with Various Bandwidths

A. Boys

		Bandwidth on Right:										
		3	4	5	6	7	8	9	10	11	12	13
Bandwidth on Left	3	0.320	0.307	0.271	0.247	0.224	0.218	0.203	0.204	0.209	0.212	-0.501
		(0.145)	(0.142)	(0.139)	(0.138)	(0.137)	(0.136)	(0.136)	(0.136)	(0.135)	(0.135)	(0.132)
	4	0.292	0.279	0.243	0.219	0.196	0.191	0.176	0.177	0.182	0.185	-0.529
		(0.112)	(0.108)	(0.105)	(0.103)	(0.101)	(0.100)	(0.100)	(0.099)	(0.099)	(0.099)	(0.095)
	5	0.293	0.280	0.244	0.220	0.197	0.192	0.177	0.178	0.183	0.186	-0.528
		(0.093)	(0.087)	(0.083)	(0.081)	(0.079)	(0.078)	(0.077)	(0.077)	(0.076)	(0.076)	(0.071)
	6	0.265	0.252	0.215	0.192	0.169	0.163	0.148	0.149	0.154	0.157	-0.556
		(0.086)	(0.080)	(0.076)	(0.073)	(0.071)	(0.070)	(0.069)	(0.068)	(0.068)	(0.067)	(0.061)
	7	0.268	0.255	0.219	0.195	0.172	0.166	0.152	0.153	0.158	0.161	-0.553
		(0.081)	(0.074)	(0.070)	(0.066)	(0.065)	(0.063)	(0.062)	(0.061)	(0.061)	(0.060)	(0.053)
	8	0.245	0.232	0.196	0.172	0.149	0.143	0.129	0.130	0.134	0.138	-0.576
		(0.077)	(0.070)	(0.065)	(0.062)	(0.059)	(0.058)	(0.057)	(0.056)	(0.055)	(0.055)	(0.047)
	9	0.212	0.199	0.163	0.139	0.116	0.111	0.096	0.097	0.102	0.105	-0.609
		(0.073)	(0.066)	(0.061)	(0.058)	(0.055)	(0.054)	(0.053)	(0.052)	(0.051)	(0.050)	(0.042)
	10	0.208	0.195	0.159	0.135	0.112	0.106	0.092	0.093	0.097	0.100	-0.613
		(0.071)	(0.064)	(0.059)	(0.055)	(0.052)	(0.051)	(0.050)	(0.049)	(0.048)	(0.047)	(0.038)
	11	0.201	0.187	0.151	0.128	0.105	0.099	0.084	0.085	0.090	0.093	-0.620
		(0.070)	(0.062)	(0.057)	(0.053)	(0.051)	(0.049)	(0.048)	(0.047)	(0.046)	(0.045)	(0.035)
	12	0.191	0.178	0.142	0.119	0.096	0.090	0.075	0.076	0.081	0.084	-0.629
		(0.069)	(0.061)	(0.055)	(0.052)	(0.049)	(0.047)	(0.046)	(0.045)	(0.044)	(0.043)	(0.033)
	13	0.175	0.162	0.126	0.103	0.080	0.074	0.059	0.060	0.065	0.068	-0.645
		(0.068)	(0.060)	(0.054)	(0.050)	(0.047)	(0.046)	(0.044)	(0.043)	(0.042)	(0.042)	(0.031)
	14	0.166	0.153	0.117	0.093	0.070	0.065	0.050	0.051	0.056	0.059	-0.655
		(0.067)	(0.059)	(0.053)	(0.049)	(0.046)	(0.045)	(0.043)	(0.042)	(0.041)	(0.040)	(0.029)

Notes: This table presents estimates from local linear regression discontinuity specifications of the effect of having an IQ score ≥ 116 on the event of completing high school on time and entering college the next year. The range of IQ's included in the sample to the left of the cutoff is indicated by the row headings; the range to the right is indicated by the column heading. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

B. Girls

	Bandwidth on Right:										
Bandwidth on Left	3	4	5	6	7	8	9	10	11	12	13
3	0.041	0.056	0.046	0.047	0.039	0.052	0.055	0.057	0.055	0.074	-0.684
	(0.141)	(0.137)	(0.134)	(0.132)	(0.131)	(0.131)	(0.130)	(0.130)	(0.130)	(0.129)	(0.127)
4	0.031	0.046	0.036	0.037	0.028	0.042	0.045	0.047	0.045	0.064	-0.694
	(0.111)	(0.105)	(0.102)	(0.100)	(0.098)	(0.098)	(0.097)	(0.096)	(0.096)	(0.096)	(0.092)
5	0.058	0.073	0.062	0.064	0.055	0.069	0.071	0.074	0.072	0.091	-0.667
	(0.095)	(0.088)	(0.083)	(0.081)	(0.079)	(0.078)	(0.077)	(0.077)	(0.076)	(0.076)	(0.071)
6	0.037	0.052	0.042	0.043	0.035	0.048	0.051	0.053	0.051	0.070	-0.688
	(0.086)	(0.078)	(0.073)	(0.070)	(0.068)	(0.067)	(0.066)	(0.066)	(0.065)	(0.065)	(0.059)
7	-0.008	0.007	-0.003	-0.002	-0.010	0.004	0.006	0.008	0.007	0.025	-0.732
	(0.082)	(0.073)	(0.068)	(0.065)	(0.063)	(0.062)	(0.060)	(0.060)	(0.059)	(0.059)	(0.052)
8	0.019	0.034	0.024	0.025	0.017	0.031	0.033	0.035	0.034	0.052	-0.705
	(0.078)	(0.070)	(0.064)	(0.060)	(0.058)	(0.057)	(0.056)	(0.055)	(0.054)	(0.054)	(0.046)
9	0.010	0.025	0.015	0.016	0.007	0.021	0.024	0.026	0.024	0.043	-0.715
	(0.076)	(0.067)	(0.060)	(0.057)	(0.055)	(0.053)	(0.052)	(0.051)	(0.050)	(0.050)	(0.041)
10	0.004	0.019	0.009	0.010	0.002	0.015	0.018	0.020	0.018	0.037	-0.720
	(0.074)	(0.064)	(0.058)	(0.054)	(0.052)	(0.050)	(0.049)	(0.048)	(0.047)	(0.047)	(0.038)
11	0.008	0.023	0.013	0.014	0.005	0.019	0.022	0.024	0.022	0.041	-0.717
	(0.073)	(0.063)	(0.057)	(0.053)	(0.050)	(0.049)	(0.047)	(0.046)	(0.045)	(0.045)	(0.035)
12	0.006	0.020	0.010	0.012	0.003	0.017	0.019	0.022	0.020	0.039	-0.719
	(0.072)	(0.062)	(0.055)	(0.051)	(0.049)	(0.047)	(0.046)	(0.045)	(0.044)	(0.043)	(0.034)
13	0.018	0.033	0.023	0.024	0.016	0.030	0.032	0.034	0.033	0.051	-0.706
	(0.071)	(0.061)	(0.054)	(0.050)	(0.047)	(0.046)	(0.044)	(0.043)	(0.042)	(0.042)	(0.031)
14	0.009	0.024	0.014	0.015	0.007	0.020	0.023	0.025	0.023	0.042	-0.715
	(0.070)	(0.060)	(0.053)	(0.049)	(0.046)	(0.045)	(0.043)	(0.042)	(0.041)	(0.041)	(0.030)

Notes: This table presents estimates from local linear regression discontinuity specifications of the effect of having an IQ score ≥ 116 on the event of completing high school on time and entering college the next year. The range of IQ's included in the sample to the left of the cutoff is indicated by the row headings; the range to the right is indicated by the column heading. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). Entries in parentheses are standard errors, clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Appendix Table A6. Mean Squared Errors of Local Linear Models with Various Symmetric Bandwidths

	(1)	(2)	(3)	(4)
	Boys		Girls	
Bandwidth	RMSE	p>F	RMSE	p>F
4	0.479	0.747	0.449	0.886
5	0.480	0.712	0.445	0.957
6	0.481	0.691	0.445	0.986
7	0.481	0.672	0.446	0.910
8	0.482	0.738	0.446	0.839
9	0.482	0.540	0.448	0.914
10	0.483	0.685	0.449	0.960
11	0.484	0.770	0.449	0.983
12	0.485	0.833	0.452	0.866
13	0.486	0.790	0.452	0.850
14	0.486	0.784	0.453	0.843
15	0.486	0.784	0.453	0.843

Notes: This table compares the performance of local linear models with symmetric bandwidths that vary between 4 and 15 IQ points. Columns 1 and 3 report the root mean squared error of the local linear regression corresponding to each bandwidth. Columns 2 and 4 report the *p*-values from *F*-tests comparing the linear model to the unrestricted alternative with a dummy for each value of IQ.

Appendix Table A7. Alternative Measures of Graduation and College Enrollment Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Grad HS On-time (Full Sample)	Grad HS On-time (Cohorts 2003-2011)	Grad HS (On-time or following year)	Grad HS On-time + Enrolled Within 1 Year	Grad HS + Enrolled Within 1 Year	Grad HS On-time + Enrolled Within 2 Years
Panel A. Boys Only						
BW 8	0.08	0.07	0.05	0.30**	0.31**	0.30*
	(0.08)	(0.09)	(0.08)	(0.12)	(0.12)	(0.11)
Obs.	1,879	1,598	1,598	1,598	1,598	1,598
BW 6	0.19	0.16	0.11	0.41*	0.40*	0.36*
	(0.12)	(0.14)	(0.12)	(0.17)	(0.17)	(0.17)
Obs.	1,451	1,242	1,242	1,242	1,242	1,242
BW 10	0.05	0.04	0.04	0.18+	0.19+	0.17+
	(0.07)	(0.08)	(0.07)	(0.10)	(0.10)	(0.10)
Obs.	2,336	1,974	1,974	1,974	1,974	1,974
Panel B. Girls Only						
BW 8	0.02	0.01	0.01	0.06	0.05	0.06
	(0.05)	(0.05)	(0.05)	(0.10)	(0.10)	(0.09)
Obs.	1,647	1,420	1,420	1,420	1,420	1,420
BW 6	-0.00	-0.02	-0.02	0.03	0.03	0.08
	(0.05)	(0.06)	(0.06)	(0.13)	(0.13)	(0.12)
Obs.	1,294	1,123	1,123	1,123	1,123	1,123
BW 10	0.02	-0.00	-0.02	0.04	0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.08)	(0.08)	(0.08)
Obs.	2,048	1,757	1,757	1,757	1,757	1,757

Notes: This table presents estimates from local linear regression discontinuity specifications. The bandwidths are IQ scores in the range 106-124 (BW 8), 108-122 (BW 6), and 104-126 (BW 10). All columns show 2SLS estimates for the effect of being identified as gifted using an indicator for having an IQ score ≥ 116 as an instrument. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). The column label at the top indicates the dependent variable in each specification. Column 1 reports results for all cohorts (2003-2012) in our analysis, while Columns 2-6 report results excluding the 2012 cohort for whom our data includes only one year beyond their "on-time" graduation year. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Appendix Table A8. Regression Discontinuity Effects and Multiple Hypothesis Testing

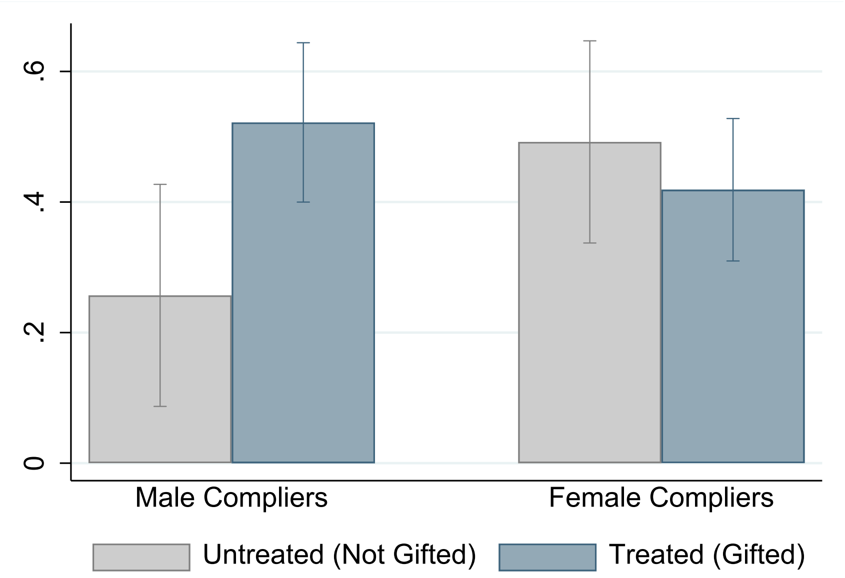
	Effect on Boys (1)	Per Comparison *p*-value (2)	Adjusted *q*-value (3)	Effect on Girls (4)	Per Comparison *p*-value (5)	Adjusted *q*-value (6)
A. Non-Cognitive Outcomes						
Gifted/High-Achieving Classroom Enrollment	0.72**	< 0.001	0.001	0.49***	< 0.001	0.001
GEM Enrollment, 6th Grade	0.26*	0.011	0.027	-0.07	0.447	0.906
G6-G8 Peers, Math Courses (σ units)	0.35**	< 0.001	0.007	0.04	0.686	0.906
G6-G8 Peers, Reading Courses (σ units)	0.48**	< 0.001	0.001	0.23***	< 0.001	0.048
Algebra Enrollment, by 8th Grade	0.30**	< 0.001	0.012	-0.04	0.715	0.906
G6-G8, No Suspensions	0.10	0.276	0.477	0.00	0.958	0.963
HS Peers, Math Courses (σ units)	0.25*	0.037	0.079	-0.02	0.840	0.939
HS Peers, Reading Courses (σ units)	0.35*	< 0.001	0.01	0.07	0.428	0.906
HS Math GPA	0.37+	0.066	0.126	-0.13	0.393	0.906
HS Lang. Arts GPA	0.11	0.564	0.825	0.01	0.963	0.963
HS GPA	0.01	0.934	0.949	-0.11	0.266	0.906
Number of AP Courses	2.47**	0.005	0.016	0.26	0.770	0.915
B. Cognitive Outcomes						
G6-G8 Test Scores (σ units)	0.04	0.814	0.949	-0.17	0.152	0.722
G6-G8 Math Test Scores (σ units)	-0.01	0.949	0.949	-0.09	0.489	0.906
G6-G8 Reading Test Scores (σ units)	0.05	0.761	0.949	-0.21	0.134	0.722
PSAT (Percentile)	0.45	0.938	0.949	-2.09	0.624	0.906
PSAT Math (Percentile)	0.60	0.919	0.949	-4.06	0.363	0.906
PSAT Reading (Percentile)	3.63	0.555	0.825	-2.79	0.565	0.906
C. College						
Grad HS On Time + Enroll Within 1 Yr.	0.28**	0.008	0.022	0.05	0.592	0.906

Notes: This table presents estimates from linear regression discontinuity specifications. The bandwidth is the set of IQ scores in the range 106-124. Columns 1 and 4 reports 2SLS estimates for the effect of being identified as gifted using an indicator for having IQ ≥ 116 as an instrument, samples of Plan B boys and girls, respectively. All estimated models exclude IQ scores that fall in the "donut" region (i.e., scores 114-116). We report standard (unadjusted) *p*-values in Columns 2 and 5. Statistical significance based on these *p*-values is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$ in Columns 1 and 4. To address concerns about multiple hypothesis testing, we use a two-step procedure from Benjamini, Krieger, and Yekutieli (2006) to calculate adjusted "*q*-values" that control for the false discovery rate (FDR), which is the proportion of rejections that are false positives (Type I errors). The adjusted *q*-values are reported in Columns 3 and 6.

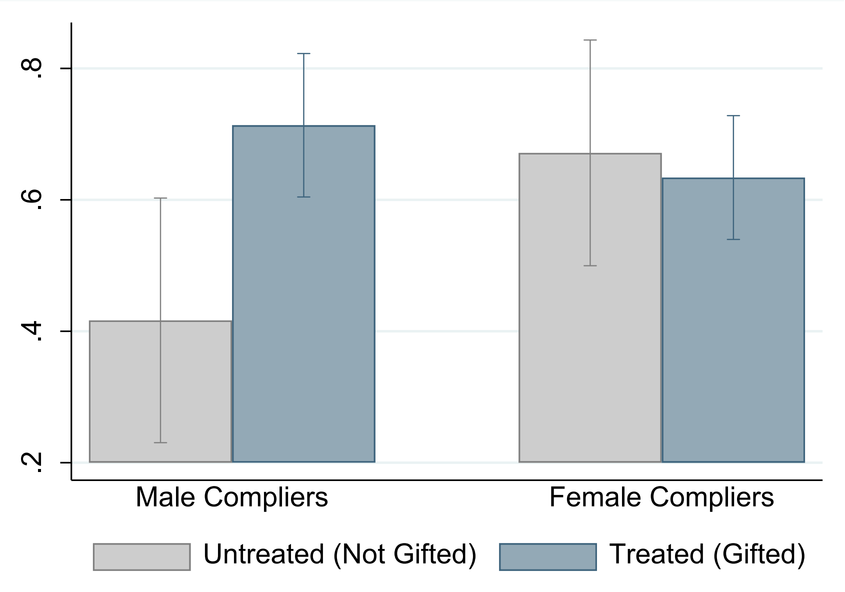
Appendix B: Mean Potential Outcomes for Complier Boys and Girls, with and without Gifted Status

Figure B1. Middle School Math Course Enrollment

A. Took Accelerated Math Track ("GEM") in 6th Grade



B. Took Algebra 1 Before High School



Notes: This figure shows mean potential outcomes for the complying boys (left side) and complying girls (right side) in our sample with and without assignment to gifted status. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.)

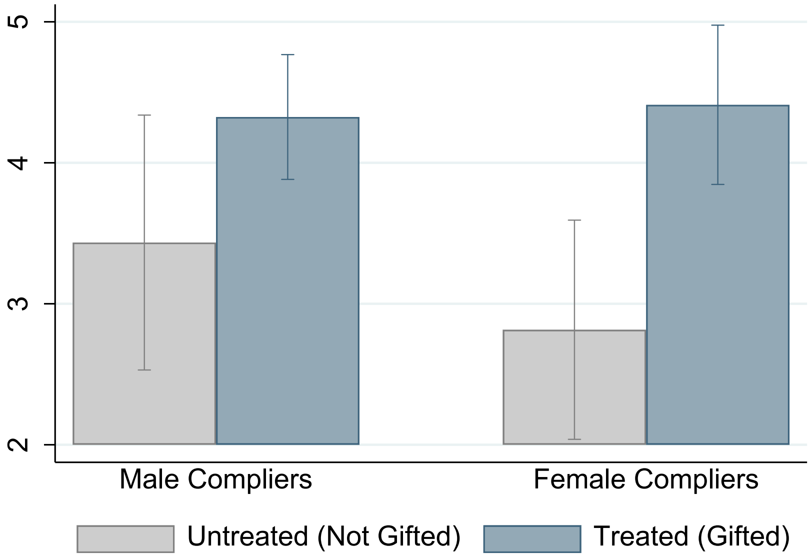
Figure B2. High School GPA's

A. High School GPA – Math Courses

B. High School GPA – Overall

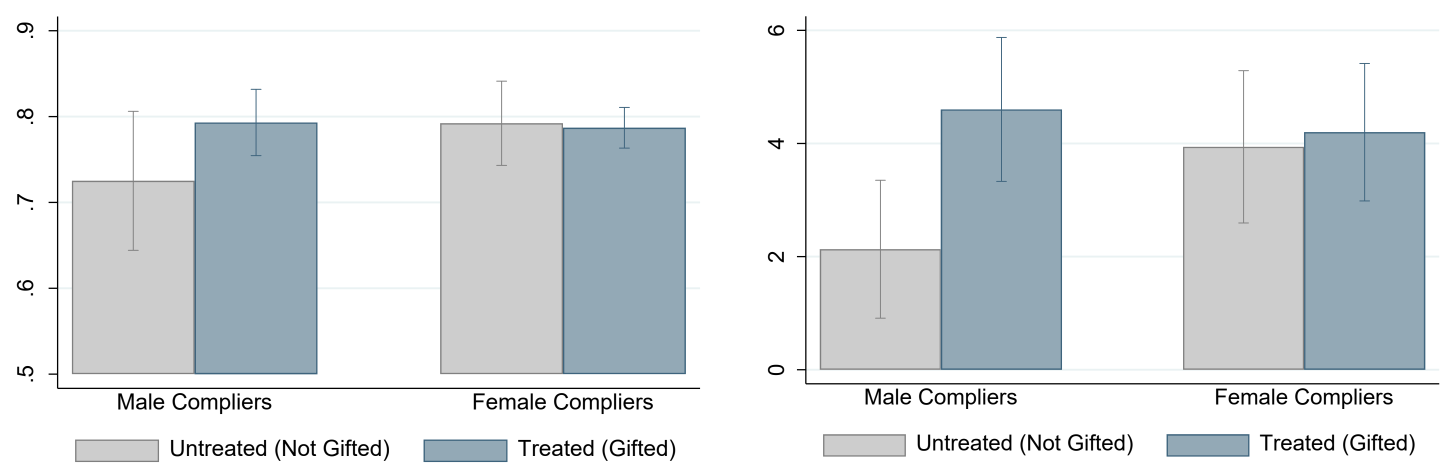
Notes: This figure shows mean potential outcomes for the complying boys (left side) and complying girls (right side) in our sample with and without assignment to gifted status. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.)

Figure B3. Survey Response on Guidance with Selecting Challenging Courses



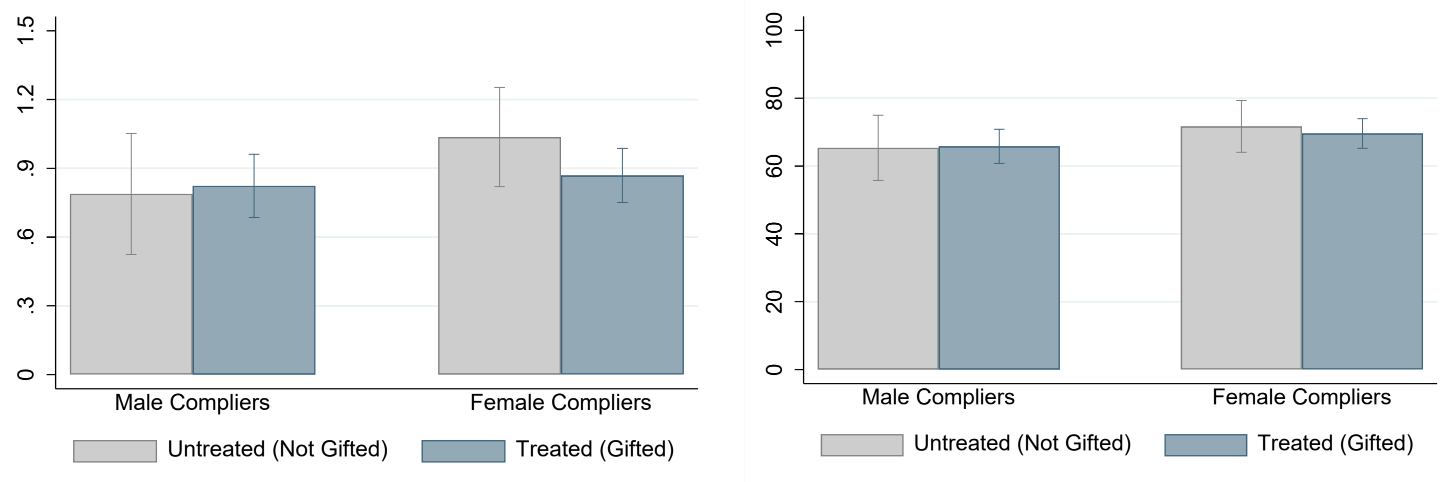
Notes: This figure shows mean potential outcomes for the complying boys (left side) and complying girls (right side) with and without assignment to gifted status. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.) The outcome is the survey response of 6th graders, collected on a 5-point Likert scale measuring agreement (from 1 = strongly disagree to 5 = strongly agree) with the statement: "This year, school staff helped me to select high level courses that challenge my abilities." The student questionnaire was only collected in the 2008-2011 5th-grade cohorts, so this analysis is based only on students in these four cohorts who had non-missing responses.

Figure B4. Additional Intermediate Outcomes



A. Grades 6-8 Extended Non-Cognitive Index

B. Number of AP Courses Taken



C. Grades 6-8 Test Scores (Reading & Math)

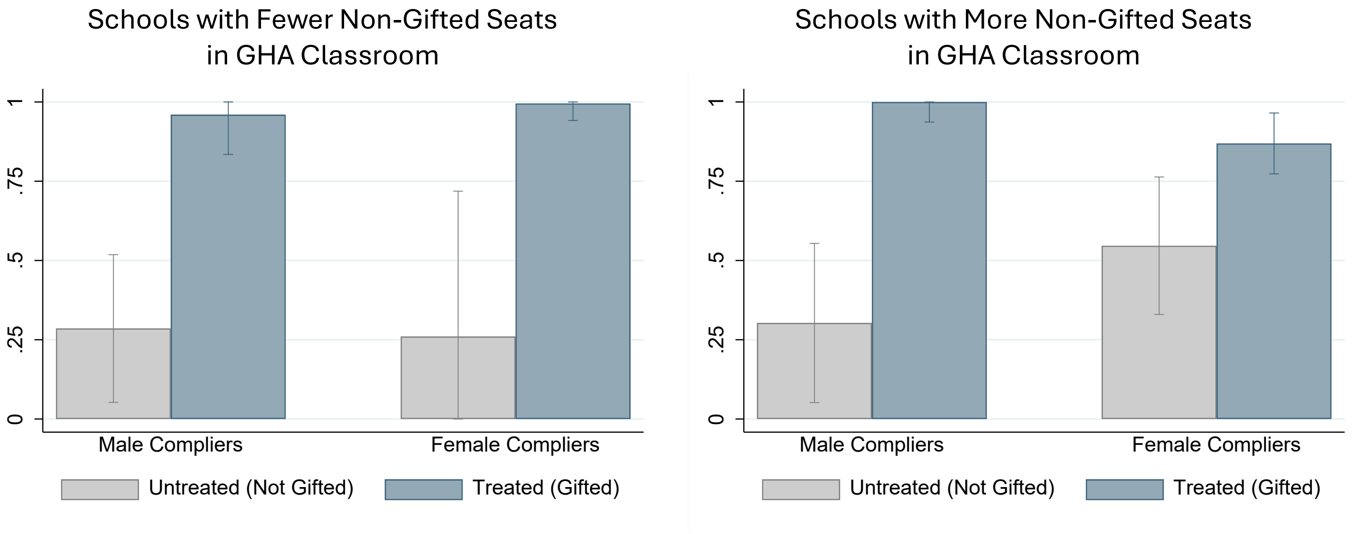
D. PSAT Score (All Subjects)

Notes: This figure shows mean potential outcomes for the complying boys (left side) and complying girls (right side) in our sample with and without assignment to gifted status. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). (See notes to Figure 6 for details.)

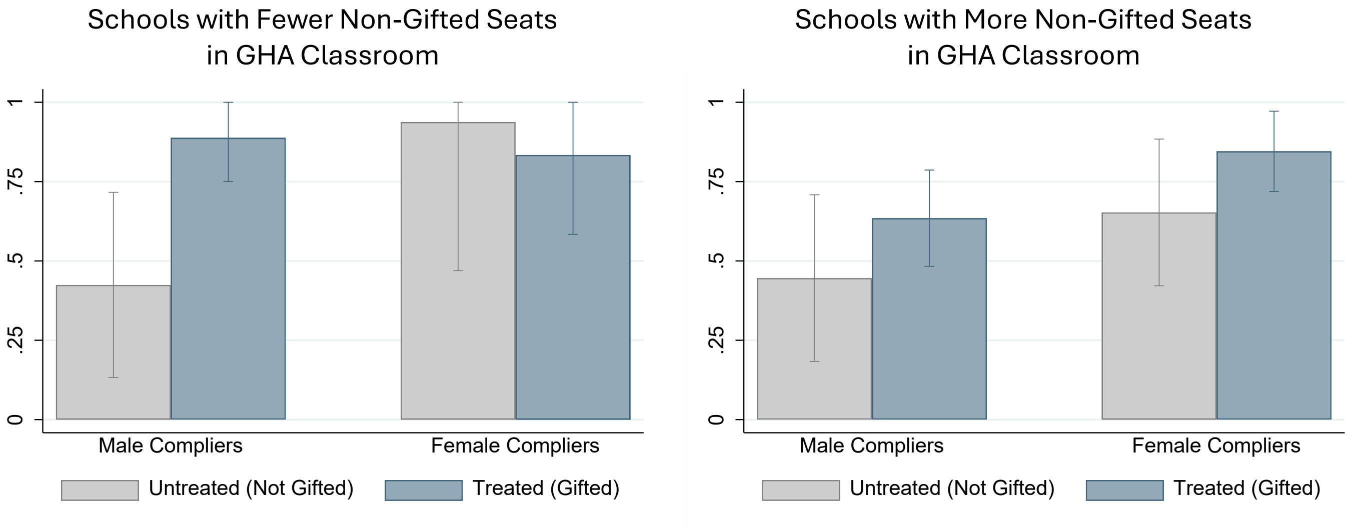
Appendix C: Analysis of Impacts of Gifted and High Achieving (GHA) Classrooms

Appendix Figure C1: Role of Participation in Gifted/High Achiever Classroom

A. GHA Participation



B. On-Time College Enrollment



Notes: This figure shows mean potential outcomes for the complying boys (left side) and complying girls (right side) in our sample with and without assignment to gifted status. The estimates are constructed by applying the approach of Abadie (2002) to local linear RD IV specifications using a symmetric bandwidth of 8 (i.e., IQ scores between 106-124) and excluding IQ scores that fall in the “donut” region (i.e., scores 114-116). (See notes to Figure 6 for details.) Panels on left use students enrolled in schools where >30% of GHA classroom seats were typically filled by gifted students; the student-weighed mean share of non-gifted students in these classrooms was around 50%. Panels on the right use schools where <30% of GHA scats were filled by gifted students; the student-weighed mean share of non-gifted students in these classrooms was more than 80%.

Appendix Table C1. Role of Participation in Gifted/High Achiever Classroom

	(1)	(2)	(3)	(4)	(5)	(6)
	GHA Classroom Participation			On-Time College Enrollment		
	Pooled	Fewer Non-Gifted Seats in GHA Classroom	More Non-Gifted Seats in GHA Classroom	Pooled	Fewer Non-Gifted Seats in GHA Classroom	More Non-Gifted Seats in GHA Classroom
Panel A. Boys Only						
BW 8	0.68**	0.67**	0.71**	0.31**	0.46**	0.19
	(0.10)	(0.14)	(0.14)	(0.12)	(0.17)	(0.15)
Obs.	1,403	499	904	1,403	499	904
Panel B. Girls Only						
BW 8	0.45**	0.74**	0.32*	0.10	-0.10	0.19
	(0.12)	(0.24)	(0.13)	(0.12)	(0.25)	(0.13)
Obs.	1,240	428	812	1,240	428	812

Notes: This table presents estimates from local linear regression discontinuity models for the effects of gifted status on GHA classroom participation (columns 1-3) and on-time college enrollment (columns 4-6). All columns show 2SLS estimates for the effect of being identified as gifted using an indicator for having IQ ≥ 116 as an instrument. All estimated models are estimated using a symmetric 8-point bandwidth (i.e., IQ scores in the range 106-124), excluding IQ scores that fall in the "donut" region (i.e., scores 114-116). The samples in columns (1) and (4) include all boys (Panel A) or girls (Panel B) for whom we have information on their school's 4th-grade GHA classroom. This excludes the first three cohorts of our main analysis sample (those in 5th grade in 2003-2005) and a small number of students who were not enrolled in the District in 4th grade or who attended a school without a GHA classroom. Columns (2) and (5) use students enrolled in schools where >30% of GHA classroom seats were typically filled by gifted students; the student-weighed mean share of non-gifted students in these classrooms was around 50%. Columns (3) and (6) use schools where <30% of GHA seats were filled by gifted students; the student-weighed mean share of non-gifted students in these classrooms is more than 80%. Standard errors are clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

Appendix Table C2. Rank-Based RD Estimates for Effects of GHA Classroom Participation for Non-Gifted, FRL/ELL High Achievers

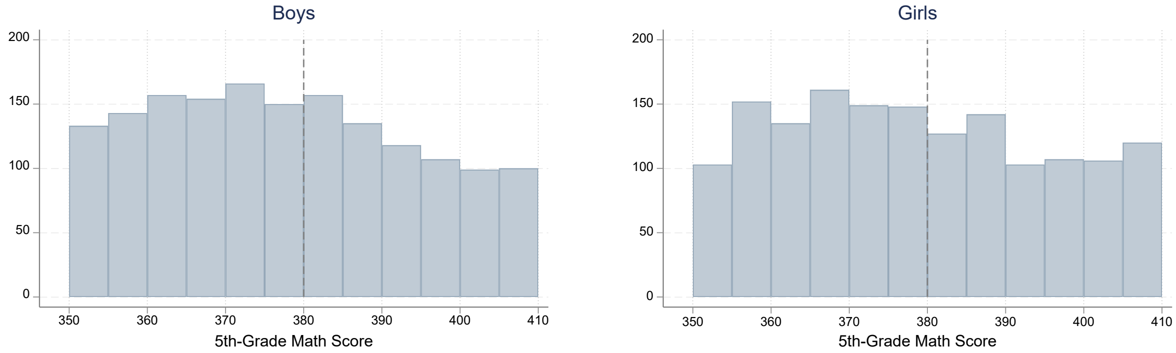
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	First Stage	**Reduced-Form**		**2SLS**						
	In G4 GHA Classroom	G3 Test Scores	High G3 Non-Cog	GEM in G6	Algebra by G8	Peer Qual. G6-G8	No Susp. G6-G8	Test Scores G6-G8	Math Scores G6-G8	Reading Scores G6-G8
Panel A. Boys Only										
BW 10 (Rank Distance)	0.20**	-0.08	0.04	0.54+	0.52+	0.30	0.68+	0.51	0.75+	0.18
	(0.06)	(0.05)	(0.07)	(0.30)	(0.27)	(0.23)	(0.38)	(0.37)	(0.43)	(0.41)
Obs.	896	896	809	896	896	896	896	894	894	894
Panel B. Girls Only										
BW 10 (Rank Distance)	0.28**	-0.02	0.01	0.09	-0.21	0.09	0.18	0.44+	0.40	0.44
	(0.06)	(0.05)	(0.07)	(0.17)	(0.14)	(0.16)	(0.19)	(0.25)	(0.28)	(0.28)
Obs.	981	981	883	981	981	981	981	975	975	975

Notes: This table presents an analysis of the effects of fourth-grade GHA participation using an alternative research design and sample based on Card and Giuliano (2016b). The running variable in the design is a student's within-school/cohort rank on an index of third-grade test scores which school officials used to determine eligibility among non-gifted children to fill any seats not taken by gifted students. Participation in GHA for non-gifted "high achievers" is instrumented using school/cohort-specific thresholds. For comparability with the Plan B gifted analysis sample, we focus on boys and girls who were eligible for the Free/Reduced Price Lunch program and English Language Learners. Of the original Card and Giuliano (2016b) analysis sample of N=4,144, there are 985 boys and 1,069 girls (2,054 children total) who meet these Plan B criteria. When we restrict the sample further to children who are observed through grade 8 with complete information on all non-cognitive outcomes in grades 6-8, we are left with a sample of 896 boys and 981 girls. All results are from linear regression discontinuity specifications. The bandwidths for the linear specifications are a student's test score rank being within 10 points of a school-specific threshold for GHA enrollment. The column label at the top indicates the dependent variable in each specification. Standard errors are clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

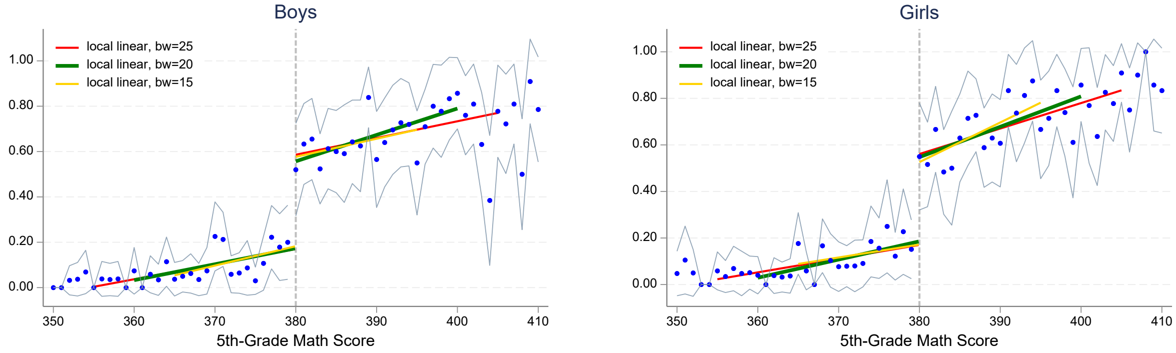
Appendix D: Regression Discontinuity Analysis of Impacts of Middle School Math Acceleration (GEM)

Appendix Figure D1: Running Variable Distribution and First-Stage Relationships

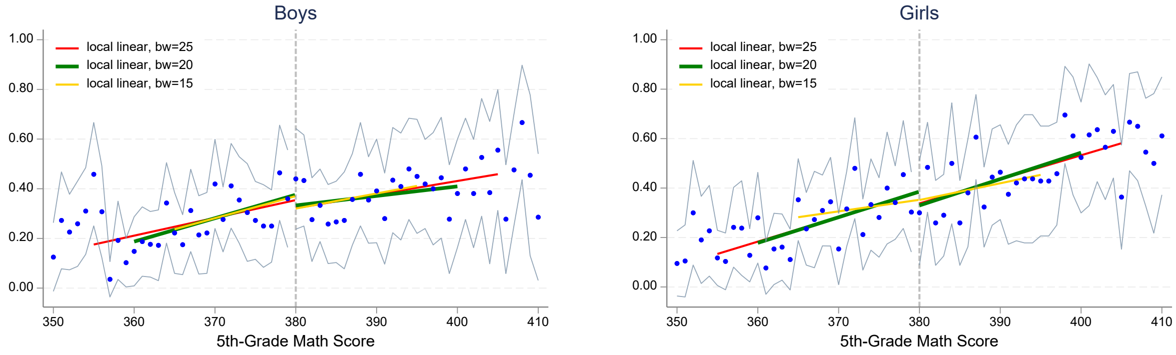
A. Histogram of Running Variable (5th Grade Standardized Math Score)



B. Participated in Accelerated Middle School Math Program (GEM) in 6th Grade

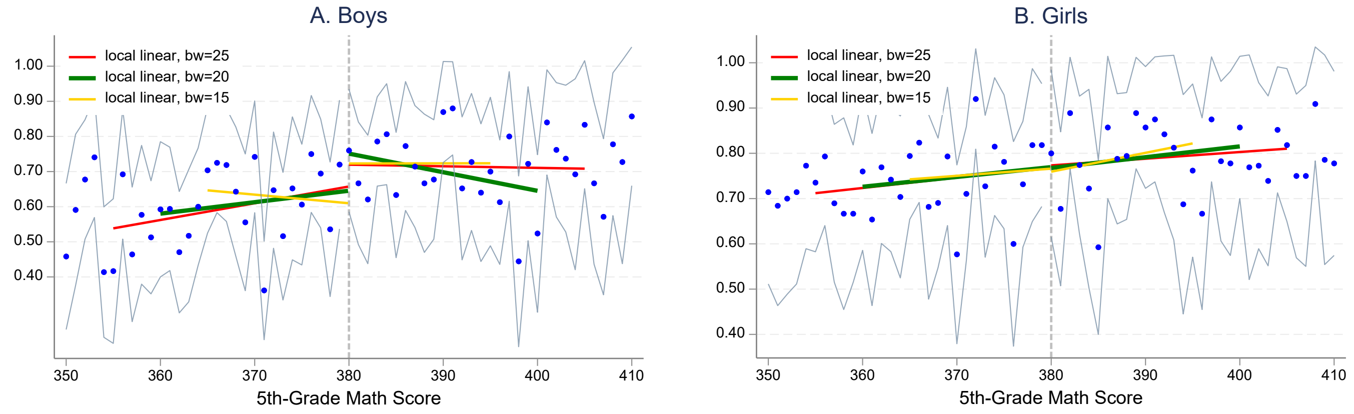


C. Identified as Gifted by 5th Grade



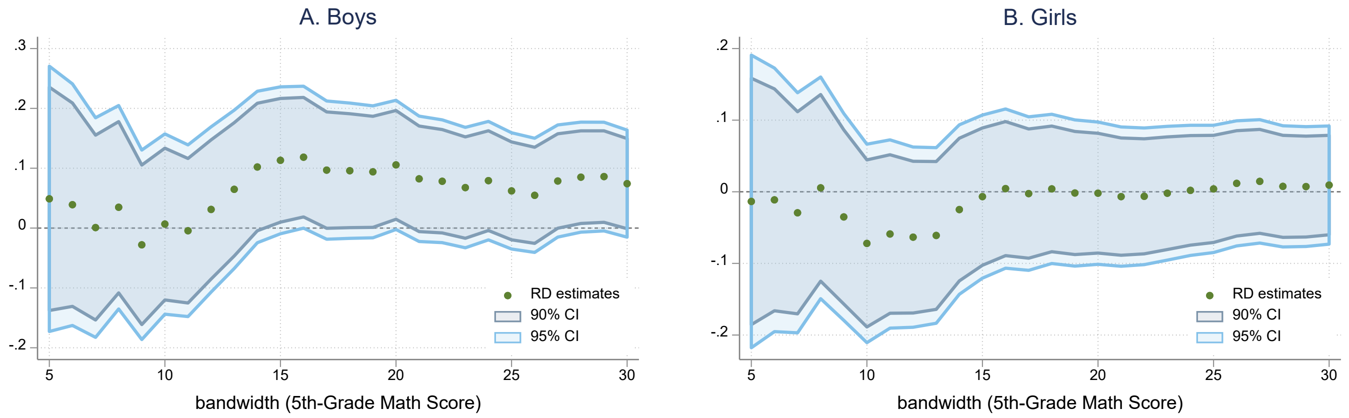
Notes: Panel A shows histograms of the student 5th-grade math score that serves as the running variable in our RD models for the effects of GEM participation. The samples include boys (left column) and girls (right column) from our IQ-based RD analysis sample whose 5th-grade math score was within +/-30 points of the GEM eligibility threshold of 380. We exclude the final cohort (i.e., those who completed 5th grade in 2012) due to a change in the scale of state-wide achievement scores in 2012. Panel B shows the first-stage relationship for GEM participation (for which district policy creates a discontinuity at the threshold score of 380), while Panel C shows an analogous plot for gifted identification (which is not expected to vary discontinuously at the GEM eligibility threshold). Panels B and C show means at each test-score value (blue dots) for students within +/-30 points of the threshold, along with 95% confidence intervals for the mean values (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 25 (red), 20 (green), or 15 (yellow).

Appendix Figure D2: Reduced-Form Relationship for Impact of GEM Eligibility on On-Time Graduation and College Entry



Notes: These figures illustrate the reduced-form relationships for boys (Panel A) and girls (Panel B) between the fraction of students who graduate high school on time and enter college and their scores on the 5th-grade math test used to determine eligibility for GEM. As in Figure D1, the samples includes boys and girls from our IQ-based RD analysis sample whose score was within +/-30 points of the GEM eligibility threshold of 380. The figures plot means (blue dots) at each test score value along with 95% confidence intervals for the mean values (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 25 (red), 20 (green), or 15 (yellow).

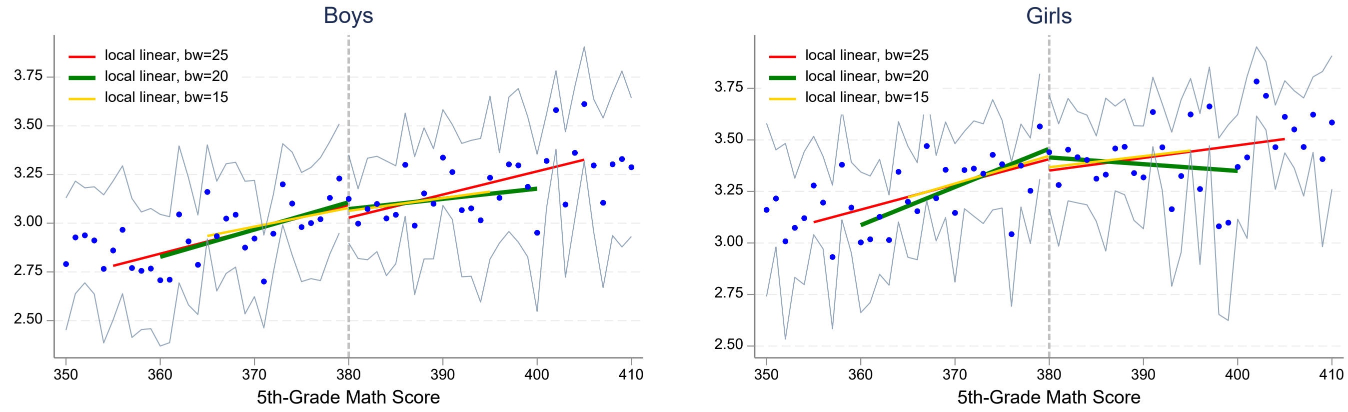
Appendix Figure D3: Reduced-Form Estimates for Impacts of GEM Eligibility on College Entry, Local Linear Models with Varying Bandwidths



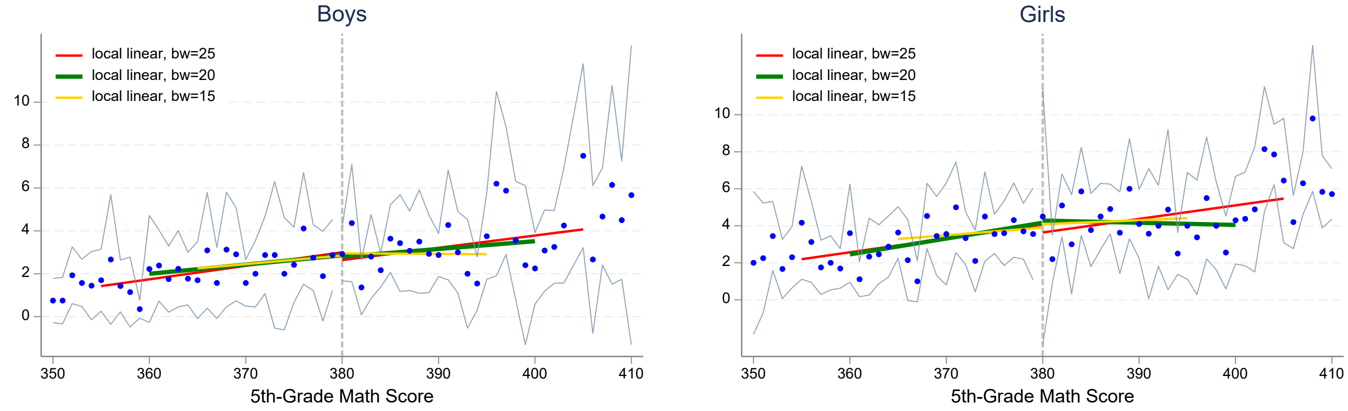
Notes: This figure plots reduced-form discontinuity estimates from local linear RD models with varying symmetric bandwidths to the left and right of the GEM eligibility threshold, separately for boys (Panel A) and girls (Panel B). The running variable is the student's score on the end-of-year standardized math test in 5th grade, and the dependent variable is an indicator for graduating from high school on time and enrolling in college the following year. The samples include all boys and girls from our IQ-based RD analysis sample whose 5th-grade math score was within +/-30 points of the GEM eligibility threshold of 380 – with the exception of the final cohort (i.e., those who completed 5th grade in 2012), who we exclude due to a change in the scale of state-wide achievement scores in 2012.

Appendix Figure D4: Impacts of Math Acceleration on High School Grades and AP Courses

A. High School GPA – Math Courses



B. Number of AP Courses



Notes: These figures illustrate the reduced-form relationships for boys (left) and girls (right) between the dependent variable (indicated by the subhead for each panel) and scores on the 5th-grade math test used to determine eligibility for GEM. As in Figure D1, the samples includes boys and girls from our IQ-based RD analysis sample whose score was within +/-30 points of the GEM eligibility threshold of 380. The figures plot means (blue dots) at each test score value along with 95% confidence intervals for the mean values (thin blue lines) and fitted values from local linear RD models with symmetric bandwidths of 25 (red), 20 (green), or 15 (yellow).

Table D1. Regression Discontinuity Effects of GEM for On-time Graduation and College Enrollment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline (Pre-gifted) Measures				First-Stage	Reduced Form	2SLS
	Grade 2, average *z*-score (Math and Reading)	Grade 3, average *z*-score (Math and Reading)	Predicted On-time Enrollment	Identified as Gifted	GEM Enrollment, 6th Grade	Grad HS On time + Enrolled Within 1 Year	Grad HS On time + Enrolled Within 1 Year
Panel A. Boys Only							
BW 20	-0.08	-0.04	0.00	-0.04	0.38**	0.11*	0.28+
	(0.09)	(0.07)	(0.02)	(0.06)	(0.05)	(0.05)	(0.14)
Obs.	935	1,135	1,165	1,165	1,165	1,165	1,165
BW 15	-0.05	0.07	0.01	-0.05	0.39**	0.11+	0.29+
	(0.11)	(0.07)	(0.02)	(0.06)	(0.06)	(0.07)	(0.17)
Obs.	718	875	900	900	900	900	900
BW 25	-0.08	-0.02	0.00	-0.03	0.42**	0.06	0.15
	(0.08)	(0.06)	(0.01)	(0.05)	(0.04)	(0.05)	(0.12)
Obs.	1,121	1,363	1,404	1,404	1,404	1,404	1,404
Panel B. Girls Only							
BW 20	0.01	-0.06	-0.02+	-0.06	0.36**	-0.00	-0.01
	(0.09)	(0.06)	(0.01)	(0.05)	(0.06)	(0.04)	(0.12)
Obs.	872	1,070	1,093	1,093	1,093	1,093	1,093
BW 15	0.06	-0.03	-0.01	-0.00	0.35**	-0.01	-0.02
	(0.11)	(0.06)	(0.01)	(0.07)	(0.07)	(0.05)	(0.14)
Obs.	681	835	851	851	851	851	851
BW 25	-0.02	-0.05	-0.02*	-0.04	0.39**	0.00	0.03
	(0.07)	(0.05)	(0.01)	(0.05)	(0.05)	(0.04)	(0.11)
Obs.	1,080	1,321	1,352	1,352	1,352	1,352	1,352

Notes: This table presents estimates from linear regression discontinuity specifications. The bandwidths for the linear specifications are 5th-grade math test scores in the range 360-400 (BW 20), 365-395 (BW 15), and 355-405 (BW 25). The column label at the top indicates the dependent variable in each specification. In columns 1-6, the entries are coefficients on an indicator for having a 5th-grade standardized math score ≥ 380. Column 7 shows 2SLS estimates for the effect of GEM enrollment in 6th grade, using an indicator for 5th-grade math score ≥ 380 is an instrument. Standard errors are clustered at the school level. Statistical significance is denoted by: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Appendix Table D2. Comparison of Plan B Gifted and GEM Compliers

	(1)	(2)
	Plan B, Gifted, IQ Scores 106-124	
	Plan B, Complier Boys	GEM RD, Complier Boys
Student Characteristics		
White	0.12	0.23
	(0.16)	(0.06)
Black	0.67	0.32
	(0.24)	(0.07)
Hispanic	0.17	0.35
	(0.21)	(0.07)
FRL	0.61	0.86
	(0.18)	(0.06)
ELL	0.14	0.06
	(0.09)	(0.03)
Parent Speaks English	0.47	0.47
	(0.20)	(0.07)
G2 Avg. Test *z*-score	0.08	0.66
	(0.54)	(0.12)
G2 Avg. Math Test z-score	0.59	0.71
	(0.46)	(0.13)
G2 Avg. Reading Test z-score	-0.35	0.6
	(0.74)	(0.15)
G3, Avg. Test *z*-score	0.77	0.83
	(0.24)	(0.09)
G3, Avg. Math *z*-score	0.78	0.86
	(0.27)	(0.09)
G3, Avg. Reading *z*-score	0.73	0.79
	(0.29)	(0.12)
G3, High Non-Cognitive Index	0.57	0.69
	(0.16)	(0.09)
Student's School Characteristics		
Share Black	0.39	0.4
	(0.11)	(0.04)
Share Hispanic	0.26	0.25
	(0.05)	(0.02)
Share FRL	0.62	0.56
	(0.11)	(0.04)
Median HH Income	45.84	53.89
	(7.90)	(2.41)
Avg. Test *z*-score	-0.08	-0.04
	(0.13)	(0.05)

Notes: This table presents estimates of the mean characteristics of compliers with the treatment defined as being enrolled in the GEM program in grade 6. Column 1 shows characteristics of the male compliers in our design for disadvantaged gifted students whose IQ scores met the gifted threshold of 116. Column 2 shows characteristics of the male compliers in the GEM eligibility RD design used in Section VI.b. Complier means are estimated by applying the approach of Abadie (2002).