Ohio’s Lost Einsteins:

The inequitable outcomes of early high achievers

Scott Imberman

September 2021

THOMAS B. FORDHAM INSTITUTE
ADVANCING EDUCATIONAL EXCELLENCE

Ohio’s Lost Einsteins:
The inequitable outcomes of early high achievers

Scott Imberman
September 2021

About the author

Dr. Scott Imberman is a professor of economics and education policy at Michigan State University. Imberman has written extensively on issues in K–12 education, including publications on school accountability, special education, and gifted and talented programs. He holds a doctorate in economics from the University of Maryland.

The Thomas B. Fordham Institute promotes educational excellence for every child in America via quality research, analysis, and commentary, as well as advocacy and charter school authorizing in Ohio. It is affiliated with the Thomas B. Fordham Foundation, and this publication is a joint project of the Foundation and the Institute. For further information, please visit our website at www.fordhaminstitute.org or write to the Institute at P.O. Box 82291, Columbus, OH 43202. The Institute is neither connected with nor sponsored by Fordham University.

Contents

Foreword

By Aaron Churchill and Michael J. Petrilli

Most agree that America must increase the racial and socioeconomic diversity of its selective high schools and colleges, as well as its top professions. Yet debate rages over how best to achieve that praiseworthy goal. In K–12 education, the impulse has often been to relax standards or even give up on notions of academic excellence or giftedness in the name of equity and inclusion. While this approach is appealing to some, it risks lowering the ceiling, holding back many young people who are ready to zoom ahead, and in the long run, diminishing the country's competitiveness.

A surer approach is to increase the supply of top-notch education offerings in order to help more high-potential kids from disadvantaged backgrounds compete successfully with others of the nation's best and brightest. That's a tall order and one that U.S. education has long struggled to achieve. Studies by Johns Hopkins professor Jonathan Plucker and others have documented stark "excellence gaps" between high- and low-income pupils. A groundbreaking analysis of patent data by Harvard's Raj Chetty and colleagues revealed that high-achieving children from low-income backgrounds rarely become inventors—what they call America's "lost Einsteins."[1]

We at the Thomas B. Fordham Institute have also sounded the alarm about a meritocratic ladder with missing rungs and too narrow a top, as well as the inadequate attention given by U.S. schools to academically talented children. In 2011, we released a study that examined a nationally representative sample of third grade high-achievers and found that a sizeable number "lost altitude" by eighth grade. More recently, we published an analysis showing that Black and Hispanic students participate in gifted programs at lower rates than their peers (a phenomenon we termed the "gifted gap"). Our team also published a book showing that while Black and Hispanic students have been participating in larger numbers in Advanced Placement classes, their passing rates on AP exams show the effects of weak preparation in elementary and middle schools.[2]

Despite much exposure of these problems and advocates' valiant efforts, our nation's most-able students continue to be overlooked—and if they come from poorer homes and neighborhoods, the overlooking is almost certain to be worse. According to the National Association for Gifted Children, twenty states don't even bother to report the number of students identified as gifted, twenty-three states fail to provide any funding for gifted education, and thirty-two states do not require universal screening for giftedness.[3] At the federal level, the sole program dedicated to supporting gifted students is the tiny Jacob K. Javits Gifted and Talented Students Education Program.

Our home state of Ohio has gotten a couple of key policies right. For one, it has long required school districts to identify students as gifted and talented using a broad definition and a variety of metrics. More recently, the state adopted a universal screening policy for giftedness. Ohio incorporates measures of achievement and growth among gifted students in its school accountability system, and it provides a modest sum of funds to support identification and services. All good, as far as they go. But are they enough?

Sensing a need to learn more, we decided to dig deep into the outcomes of high-achieving students in Ohio. How are smart kids there doing? Are they staying on top from elementary through high school? Are most of them

---

1 Jonathan A. Plucker, Nathan Burroughs, and Ruiting Song, *Mind the (Other) Gap! The growing excellence gap in K-12 education*, Center for Evaluation and Education Policy (2010); Joshua S. Wyner, John M. Bridgeland, and John J. Diulio, *Achievement Trap: How America is failing millions of high-achieving students from lower-income families*, Jack Kent Cooke Foundation (2007); and Alex Bell, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John van Reenan, "Who Becomes an Inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics* (2019).

2 Yun Xiang, Michael Dahlin, John Cronin, Robert Theaker, and Sarah Durant, *Do High Flyers Maintain their Altitude? Performance trends of top students*, Thomas B. Fordham Institute (2011); Christopher Yaluma and Adam Tyner, *Is There a Gifted Gap? Gifted education in high-poverty schools*, Thomas B. Fordham Institute (2018); and Chester E. Finn, Jr., and Andrew E. Scanlon, *The Role of Advanced Placement in Bridging Excellence Gaps*, Thomas B. Fordham Institute (2020).

3 Anne N. Rinn, Rachel U. Mun, and Jaret Hodges, *2018-2019 State of the States in Gifted Education*, National Association for Gifted Children and the Council of State Directors of Programs for the Gifted (2020).

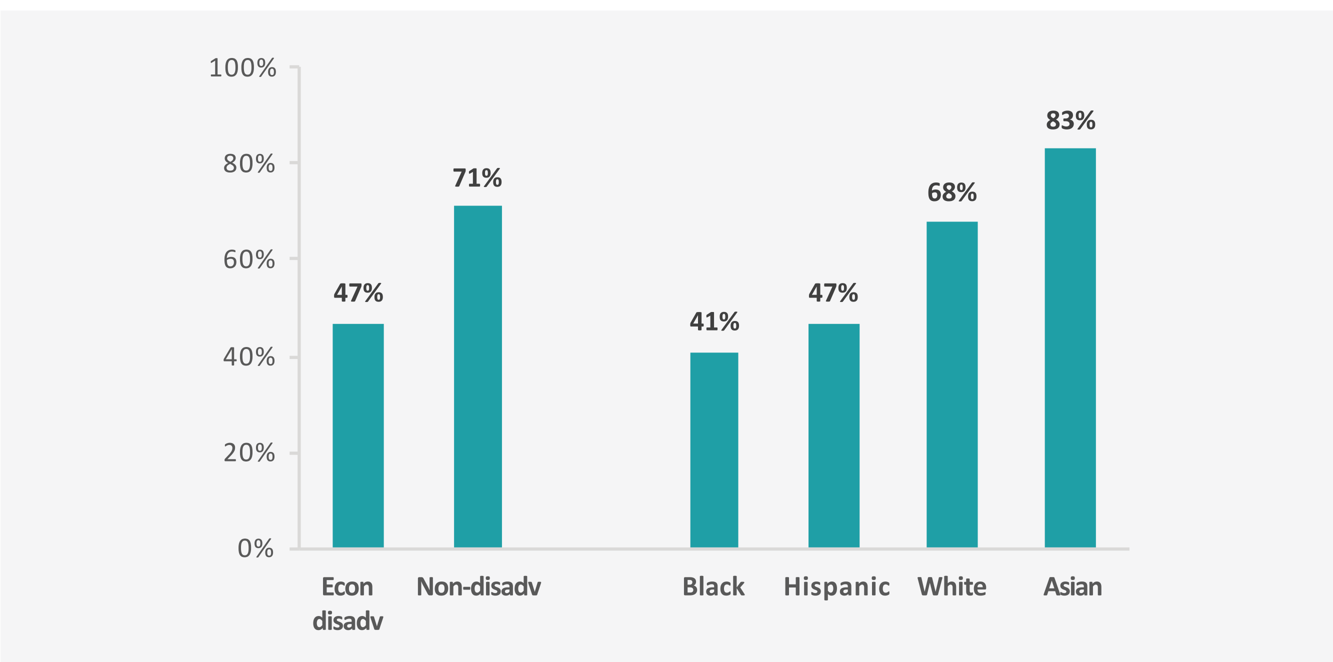
acing their college exams and heading to four-year universities? Do we see disparities by race and socioeconomic status, even among students who do well on their third grade exams? And what about being formally identified as gifted—does that provide an extra boost?

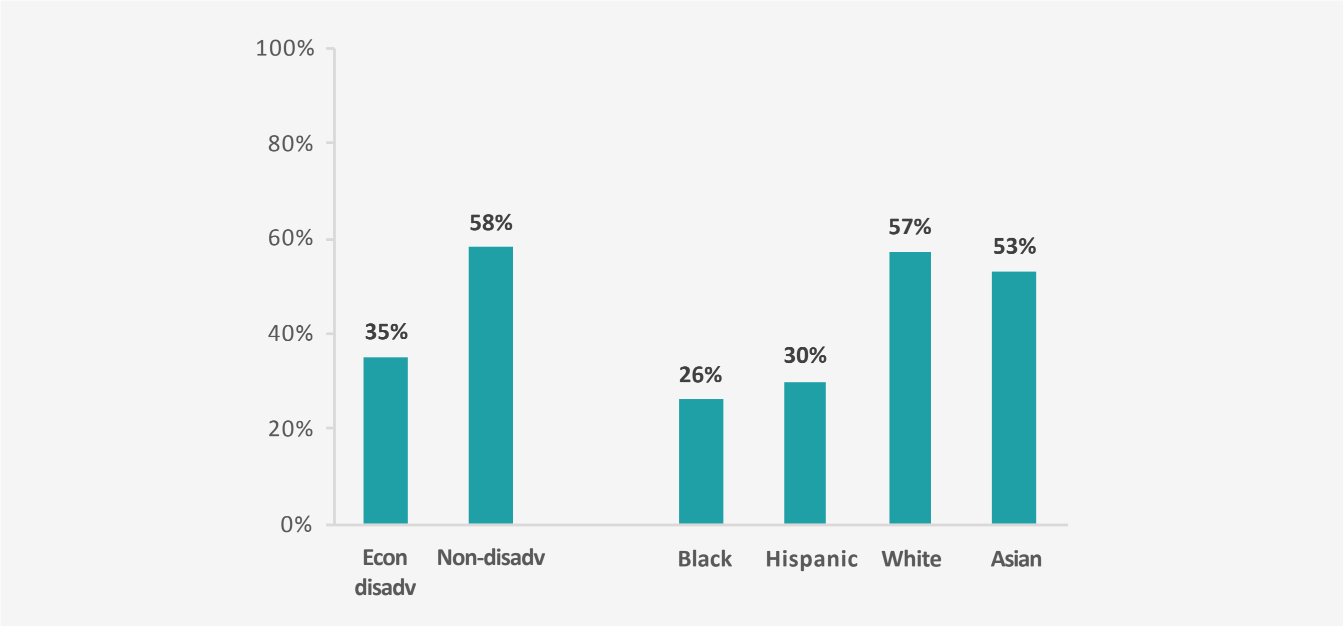
To examine these questions, we turned to Scott Imberman of Michigan State University, a first-rate economist whose work includes a widely cited study on the effects of gifted programs. We're pleased to present his comprehensive study, which relies on Ohio's longitudinal databases containing anonymous student-level records. The focus of his analysis is the academic trajectory of the state's "early high achievers," defined as students who scored in the top 20 percent statewide on third grade exams in math or English language arts (ELA).

The first part of the report documents the basic outcomes of these early high achievers. Regrettably—but consistent with previous national studies—we see "excellence gaps" emerge by economically disadvantaged status and for Black students.

- On grades 4–8 state exams, economically disadvantaged early high achievers lost ground to their nondisadvantaged, high-achieving peers, indicating that excellence gaps tend to widen over time. In math, for example, disadvantaged high achievers' scores fell by 0.18 standard deviations, while their nondisadvantaged peers lost 0.11 SDs between grades 4 and 8. In a similar vein, Black early high achievers made less progress over time on state tests than their peers from other races.
- Less than half of economically disadvantaged and Black early high achievers took the ACT—47 and 41 percent, respectively—which is the predominant college entrance exam used in Ohio. This compares with 71 percent of non-disadvantaged high achievers who participated in this test. (During the period of this study, ACT or SAT participation was voluntary in Ohio.)
- The average ACT math and reading scores of economically disadvantaged and Black early high achievers fell short of their more advantaged peers. For instance, the average ACT math score for economically disadvantaged high achievers was twenty-three, while the average score for non-disadvantaged high achievers was twenty-five.
- Just 35 percent of economically disadvantaged and 26 percent of Black early high achievers went on to enroll in four-year colleges. This compares with 58 percent of non-economically disadvantaged high achievers who enrolled in such institutions.

Figure A: ACT test participation (top) and 4-year college enrollment (bottom) among early high achievers

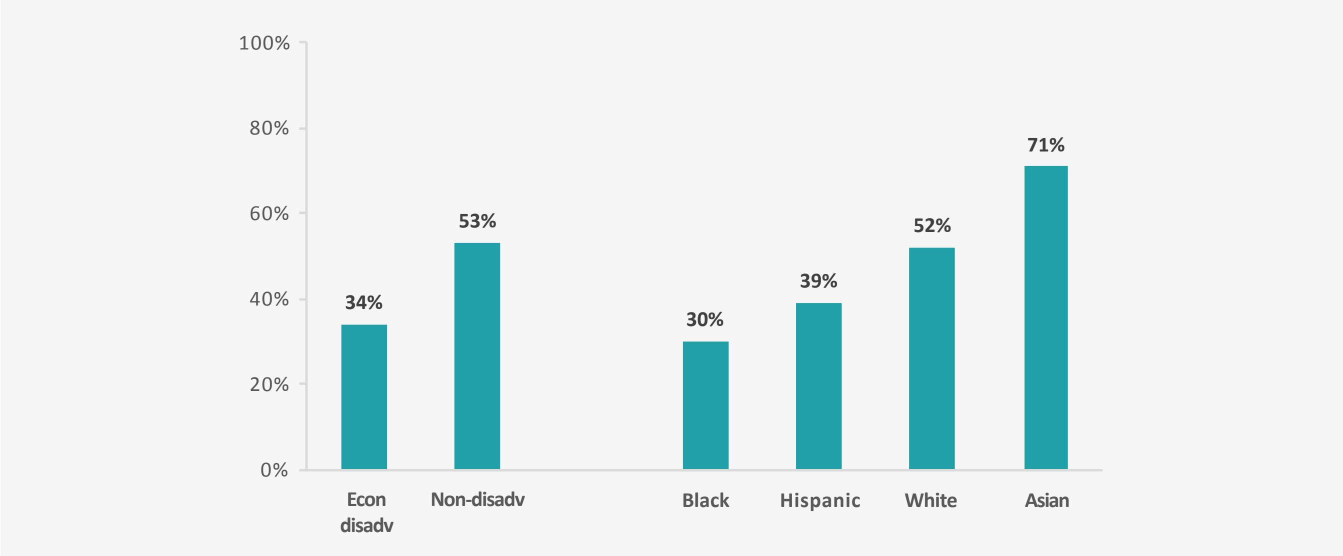




The report then turns to questions of gifted identification. Here we see that fewer than half of Ohio's early high achievers are formally identified as gifted by eighth grade, reflecting either a high bar for identification or perhaps uneven application of the criteria for identification.[4] Thus, the terms "early high achiever" and "gifted student" are not interchangeable—and the raw outcomes data show that gifted high achievers outperform their non-identified peers.

In terms of who's identified as gifted among early high achievers, economically disadvantaged and Black students were less likely to be identified. By eighth grade, 34 percent of economically disadvantaged high achievers were identified, compared to 53 percent of their nondisadvantaged peers. Meanwhile, 30 percent of Black early high achievers were identified as gifted versus 52 percent of White and 71 percent of Asian high achievers. Note that, during the period of this analysis, Ohio did not have a universal screening policy—that changed in 2017—possibly explaining some of the disparity in identification rates.

Figure B: Gifted identification by grade 8 among early high achievers



4 To be identified as gifted in math, students must score at the ninety-fifth percentile or higher on an approved nationally normed test; the same bar applies for identification in reading, science, and social studies. Slightly different criteria apply for identification in superior cognitive ability, visual or performing arts, and creative thinking, the remaining categories of giftedness in Ohio. For more on Ohio's gifted identification policies and practices, see Ohio Department of Education, *Assessments Approved for Gifted Identification and Prescreening* (2021).

Rigorous causal analyses indicate that gifted identification itself provides a small boost to early high achievers from all backgrounds on state math and ELA exams, but the gains are more substantial for Black students, particularly in math. Though not causal evidence, Black high achievers who are identified as gifted outperform non-identified Black high achievers on ACT and AP performance and college-going outcomes.

Due to data limitations, this study cannot answer questions about whether *receiving* gifted services drove the results—Ohio does not mandate services, just identification—or even which specific types of service benefitted children most. The state does not produce systemic data about the services that students receive—if indeed they receive them—so we don't know whether gifted students are mostly receiving meager hour-a-week "enrichment" or more intensive programming with specialized curricula and classrooms. What happens *after* a student is identified remains hidden inside a black box.

Nevertheless, it's clear from the research—this study included—that more needs to be done on behalf of America's high achieving kids, especially those from low-income backgrounds. What to do? Education groups and policymakers can also take concrete steps to place the needs of high-ability kids on the policy agenda. Consider just a few starter ideas—some of which Ohio has already implemented (an * indicates such a policy).

- Screen all children at least once in elementary school for academic giftedness,* ideally using the state assessment rather than an additional exam.
- Ensure that high-achieving (but not formally gifted) pupils have access to gifted programs, should capacity exist.
- Include data about gifted students' achievement and growth on annual school report cards.*
- Require schools to provide gifted services, and report outcomes by the type of service received.
- Identify all students in the state who show potential for success in selective colleges by requiring high schoolers to take the ACT or SAT at least once.*
- Remove barriers to AP or IB exams by fully covering test fees for low- and middle-income students.*
- Ensure that all students are well-versed in higher education opportunities, including information about financial aid and which colleges might be an appropriate match.
- Empower parents of high-achieving children with options if their local school doesn't offer satisfactory gifted programs or advanced coursework.
- Create specialized schools that cater specifically to the needs of gifted students and high achievers.*

***

Thousands of early high-achieving children—including smart kids of poor and working-class parents from places like Cincinnati, Dayton, or Mansfield, Ohio—are going adrift as they make their way through middle and high school. They don't end up taking AP exams, achieving high marks on their ACTs, or going to four-year colleges, as one might expect. This not only limits these kids' opportunities to move up the social ladder, but also threatens the nation's economic competitiveness and derails our aspirations for a more just society where children from all backgrounds can become inventors, doctors, and business leaders. Rather than buying into the false assumption that high-achieving kids will do fine on their own, we need to do a better job making sure that all high achievers, including those from low-income backgrounds, get the education they deserve.

Acknowledgments

We extend our deepest thanks to Scott Imberman for producing this fine report. We offer special thanks to the Ohio Department of Education and CHRR at The Ohio State University for providing the data and technical support that made this project possible. We also appreciate the insightful comments from Vladimir Kogan and Stéphane Lavertu of The Ohio State University and Ann Sheldon of the Ohio Association for Gifted Children. On the Fordham team, we're grateful to Chester E. Finn, Jr., Chad L. Aldis, Amber Northern, Brandon Wright, and Adam Tyner who offered thoughtful feedback during the drafting process. We wish to thank Victoria McDougald, Olivia Piontek, Pedro Enamorado, and Jeff Murray for their support in report production and dissemination. We thank Pamela Tatz for copyediting the report and Andy Kittles for developing the layout and design. Last, we acknowledge the work of Raj Chetty and colleagues whose study on inventors helped to inspire the title of this publication.

Frequently used acronyms

AP	Advanced Placement
API	Asian and Pacific islander
ELA	English language arts
GT	Gifted and talented
IPEDS	Integrated Postsecondary Education Survey
NSC	National Student Clearinghouse
OAA	Ohio Achievement Assessment
OGT	Ohio Graduation Test
OLDA	Ohio Longitudinal Data Archive
OLS	Ordinary least squares
SD	Standard deviation

Executive summary

School accountability systems have led administrators and educators to focus on improving academic outcomes for low-performing students. There is concern that some of this focus has come at the expense of their high-achieving peers. There is also evidence that high-performing students from underrepresented minorities are less likely to have access to gifted and talented (GT) services, and rigorous studies indicate that they are more likely to benefit from them. Given concerns about the lack of diversity and differential outcomes in top colleges and universities and highly skilled occupations, as well as the potential for these institutions to be drivers of increased intergenerational social mobility, it's vital that academically able children from disadvantaged backgrounds reach their full potential.

This study examines the academic and college-going outcomes of Ohio's high-achieving students. Do they "maintain their altitude"—keep their status as high achievers—throughout middle and high school? How do they fare on college-entrance (ACT or SAT) and Advanced Placement (AP) exams? How many go to four-year colleges? Does being formally identified as gifted make a difference in outcomes? And, perhaps most importantly, are there gaps by race/ethnicity or income?

To shed light on these questions, this analysis relies on anonymous, student-level data to track "early high achievers"—children who scored in the top 20 percent on Ohio's third-grade tests—from fourth grade into college. It examines early high achievers' state exam scores, college-entrance and AP exam results, and college enrollments. And because just under half of early high achievers acquire gifted status between third and eighth grade, this study is able to use this change to gauge whether gifted identification causes improved performance on state exams.

The analyses yield the following findings on the achievement growth of early high achievers:

1. Between fourth and eighth grade, early high achievers see their achievement fall slightly or remain flat, while the performance of early low achievers increases. This is not just because of "regression to the mean." Ohio schools seem to be doing a better job of boosting the achievement of low performers over time than high performers, a pattern that has been identified in national studies, as well.
2. Achievement growth among high achievers is lower for economically disadvantaged, female, Black, and Hispanic students. Growth is higher for Asian/Pacific Islander (API) students.

The analyses also find the following on the long-run outcomes of early high achievers:

3. During high school, gaps by race and income among Ohio high achievers are seen on state standardized, college-entrance, and AP exams. Black early high achievers are the lowest performers in high school, followed by Hispanic, White, and API early high achievers, in that order. There are no sizable gender gaps along these metrics, though girls are more likely to take SAT and ACT exams.
4. Similar gaps by race and income emerge when considering the college enrollment of early high achievers. The notable exception is API high achievers, who—despite outpacing other racial groups on achievement metrics—have the lowest enrollment rates in two-year schools among early high achievers and lag White students in enrolling in four-year colleges. And among early high achievers, female enrollment in both types of colleges exceed male enrollment.
5. These gaps in achievement and college-enrollment outcomes hold even after accounting for high-achieving students' precise third-grade achievement levels, elementary school attended, and cohort. The only exception is college attendance, where API students fare worse than other races and other races enroll at similar rates.

Turning to a narrower focus on early high achievers who are identified as gifted in elementary or middle school, we find the following:

6. High achievers who are economically disadvantaged, Black, or Hispanic are substantially less likely to be identified for gifted services—which is clear evidence of GT identification gap.
7. Only 23 percent of third graders who are high achievers in math in Ohio are identified for GT services in math in third grade. By eighth grade, this only increases to 36 percent. Among early high achievers in reading, 20 percent are identified as gifted in reading in third grade and 28 percent by eighth grade.
8. High achievers who are identified for GT programs exhibit higher performance on high school achievement exams, even when compared to virtually identical students in the same school who are not so identified—by 0.4 standard deviations (SDs) in math and 0.3 SDs in ELA. These are very large differences.
9. High achievers who are identified as GT are more likely to take a college-entrance exam and to score higher on it. They are also more likely to take AP exams and score higher. These results hold after controlling for early achievement, being in the same elementary school, race, gender, and economic status. As for college enrollment, high achievers who are identified as gifted are substantially more likely to attend four-year colleges, while identified students are only slightly more likely to attend a two-year college.

Though findings 6 through 9 reveal gaps between early achievers who are identified as gifted and those who are not, those differences may not be caused by GT identification or the provision of GT services and might instead simply reflect other factors related to both GT identification and outcomes—for example, family income or other unobserved abilities such as visual arts skills. Unfortunately, due to data limitations, this study cannot uncover the impacts of GT programs or services per se. However, the data do permit an analysis of the causal impacts of being *identified* as gifted on state exam achievement. The results are as follows:

10. Being identified as gifted in math leads to a 0.03 SD increase in annual math scores of early high achievers and a 0.02 SD increase in reading scores (equivalent to gains of approximately one percentile per year). Being identified as gifted in cognitive skills leads to a 0.02 SD increase in annual math scores. No other type of identification (e.g., social studies or visual or performing arts) generates a statistically significant effect on later achievement. Note that, in Ohio, identification as gifted does not necessarily mean that a student receives GT services, so these results may underestimate the impact of actually participating in gifted programs.
11. Among early high achievers, gifted identification in math generates larger increases in math achievement for Black and Hispanic students than White pupils. Identification in reading improves reading scores for boys but not girls; no other student group demonstrates significant gains in reading. And schools with higher Black and Hispanic populations show larger increases in both math and reading achievement from GT identification.

Overall, this study provides some evidence that closing the "GT identification gap" could help to close the "excellence gap." In other words, if early high-achieving, low-income students and students of color were identified for GT programs at the same rate as their White and Asian high-achieving peers, this could reduce achievement gaps down the road, though it is by no means a panacea, as the estimated effects are relatively small. Though the evidence provided here on long-run outcomes is not causal, when combined with the achievement results, it suggests that if these third-grade "high flyers" were to gain GT identification, they would be more likely to go on to demonstrate higher achievement on AP exams and the ACT or SAT—and, likely as a result, matriculate at higher rates into four-year colleges, including elite institutions.

Introduction

Accountability systems in education that focus on student proficiency have led to increases in achievement (Hanushek and Raymond 2004; Jacob 2005; Dee and Jacob 2011). However, there's some evidence that accountability has induced schools and teachers to focus on students who are close to proficiency targets, which leads to smaller gains at higher parts of the achievement distribution (Neal and Schanzenbach, 2007).

Though acknowledging the moral imperative to prioritize low-performing students, we also need to ask whether this is costly for high-performing students—and, if so, what might reduce these costs. One much-pursued strategy is the use of GT programs. Evidence on GT's impact (and that of other advanced academic programs) has been mixed. Some studies show no effect (Abdulkadiroğlu, Angrist, and Pathak 2014; Bui, Craig, and Imberman 2014; Dobbie and Fryer 2014). But, though not always the case (Redding and Grissom 2021), studies show large positive effects for students from minority groups (Card and Giuliano 2016a; Cohodes 2020). Further, there is evidence that students from underrepresented minority groups are identified for GT at considerably lower rates than similar White students (Grissom and Redding 2016; Ford 1998). Nonetheless, it is important to note that all these studies are done in different contexts in terms of services provided and definitions of giftedness. Hence, though they give us some idea of whether gifted schooling "works" on average, the types of services that work better need more investigation.

In this study, I start with the following research question: How do high achieving—based on third-grade exam scores—Ohio students from different backgrounds perform as they progress in their education, and are there educational outcome gaps by race, economic status, gender, and gifted identification among these students? A considerable worry among educators and policymakers in Ohio and elsewhere is that the ability to maintain high performance could vary considerably by student demographics. It has been well established that achievement gaps exist between racial and economic groups, with the more advantaged groups having higher performance. Though these gaps tend to shrink throughout students' time in school, they remain into adulthood (Braun, Chapman, and Vezzu 2010; Burchinal et al. 2011; Fryer and Levitt 2004; Hanushek and Rivkin 2009; Jencks and Phillips 2011; Reardon and Galindo 2009; Reardon, Robinson-Cimpian, and Weathers 2008).

Nonetheless, we know far less about how these gaps develop when we focus on students who already exhibit high achievement early in their education. Although policymakers often justifiably focus their efforts on bringing low achievers up to higher levels, it is important to understand whether this is complementary to or a substitute for maintaining high achievers' performance. This is particularly the case for high achievers from disadvantaged backgrounds, who may be more at risk of falling behind, potentially contributing to a lack of diversity in elite schools, colleges, and the upper part of the income distribution.

I find that even when restricting to the top quintile of achievers in third grade, controlling for precise measures of third-grade achievement, and controlling for whether students are in the same school in third grade, wide gaps exist in long-term outcomes by race and economic disadvantage. In other words, high-achieving White and affluent students are much more likely to remain high achieving than their Black and lower-income peers. These descriptive analyses provide a broad picture of the educational status of high-achieving students in the Buckeye State and pinpoint student types where improvement may be needed.

I then turn to the potential role of GT in driving these trends. Though ideally one would like to consider actual service provision, limitations in the data lead me to focus the analysis on the just the *identification* of students for gifted services. I begin by showing relationships between students' gifted status and their achievement trends and long-run outcomes. Nonetheless, as with the first part of the study, this is descriptive and does not provide information on the causal impact of GT programs. I therefore turn to estimating the causal impacts of gifted identification on contemporaneous achievement and student absences, using a "fixed-effects" econometric strategy. This method allows us to remove statistical bias that is the result of unobserved characteristics of students that do not change over time.

The rest of this report progresses as follows. First, I briefly describe the methods for the study. Next, I look at trends in achievement for early high achievers as they progress through elementary and middle school. Then, I look at how early high achievers from different demographic groups compare to each other in high school outcomes and college attendance. The next sections turn the focus to early high achievers who are identified as gifted. I first briefly describe GT programs in Ohio. Then, I turn to looking at achievement trends by early gifted status and investigate long-term outcomes for high achievers who are ever identified as gifted and those who are never identified. Last, I estimate the impacts of GT identification on achievement and absences. The conclusion recaps the study, and the appendixes offer technical information and supplemental data.

Methods

This study uses three methods of assessing how high achieving students' academic performance evolves over their schooling, how outcomes differ by demographics and gifted status, and what the impact of gifted identification is on student performance.

Comparisons of means and distributions

The first method is a simple comparison of means or distributions across groups and over time. I start with simply tracking mean (average) achievement by grade level across different groups. Then, for high school and college-going outcomes for variables that are binary (e.g., they can only be zero or one), I provide means of outcomes for each demographic group or by gifted status. For variables that are continuous (can have many different values), I provide the interquartile range, which shows the twenty-fifth-percentile, mean, and seventy-fifth-percentile value.

Linear regression analysis

Following the means and distributions, I provide analyses that use ordinary least-squares (linear) regression models. These models "control" for a set of variables that provide estimates of how outcomes differ by race, gender, and economic status when comparing students with the same third-grade achievement, who entered third grade in the same cohort, and who attended the same school in third grade. More details on specific modeling are provided in Appendix A.

Student fixed effects

Although the prior two methods only provide correlations rather than causal impacts, I use the panel (over time and individuals) nature of the Ohio data to provide plausibly causal estimates of the effect of GT identification on achievement using student fixed-effects regressions. The basic idea of the model is that having multiple observations on students over time as they gain GT identification allows one to account for any unobserved differences between students that do not vary over time. Essentially, the model compares student outcomes before and after they are identified to pick up changes in trends. Hence, various innate characteristics of students (e.g., ability) are accounted for in the regression model. If one is willing to assume that the unobserved aspects of students that affect both GT identification and achievement are only made up of these "time-invariant" characteristics, then the estimates of the relationship between GT identification and achievement can be considered causal. Thus, under these reasonable assumptions, I can show the direct impact of GT identification on student performance in elementary and middle school. For a more detailed description of this empirical method, please see Appendix A.

Achievement trends for early high achievers

Research question: What do achievement growth patterns look like for students who were high achieving in third grade?

This study first considers the subsequent achievement patterns of Ohio students who were high achievers in third grade. For this analysis, I use data from the Ohio Longitudinal Data Archive (OLDA) at Ohio State University. The data include every student in Ohio who took the state math or reading exams in grade 3 between the 2005–06 and 2011–12 academic years and follows them through grade 8.[1] In total, over 900,000 students are included. Additional analyses go on to consider their outcomes in high school and college, though with fewer cohorts.

One difficulty is defining what we mean by "high achievers." Any such definition is somewhat subjective. For the purposes of this study, I identify a high achiever as being a student who scores in the top quintile (20 percent) statewide in third grade on the Ohio Achievement Assessment (OAA) exam in math (for math outcomes) or reading (for reading/English language arts [ELA] outcomes).[2] Although differentiated instruction like GT may not be appropriate for all students in this relatively wide band, I wanted to make sure that the data captured a relatively wide pool of potential GT students without expanding the sample to include too many students who would not qualify under most circumstances. This threshold, in my view, strikes this balance but is by no means definitive. Indeed, note that gifted status is typically identified based on being in the top 5 percent of a nationally normed exam, and therefore, GT students would only be expected to be a portion of the high achievers identified in this fashion. That amounts to almost 180,000 students, whose achievement is tracked through eighth grade.

<u>**Characteristics of Ohio students, by third-grade achievement**</u>

Finding 1: High achievers are less likely to be Black, Hispanic, or disadvantaged students. They also show lower growth than low achievers, on average, in reading but not math.

Who are these high-achieving students? In Appendix A, Table A1 provides means and SDs for select characteristics overall and by quintiles of third-grade achievement in math. A key pattern to note is that as achievement levels rise, students are more likely to be White and less likely to be Black or Hispanic. The results for Black students are particularly striking, though also typical of racial achievement gaps nationally. For example, despite making up 16 percent of third graders in Ohio between 2006 and 2012, they account for 34 percent of students scoring in the bottom and only 5 percent of students scoring in the top 20 percent. In addition to the racial differences, only 1.5 percent of students with limited English proficiency are high achievers, despite being 2.3 percent of the total population in Ohio, and economically disadvantaged students (those who qualify for subsidized lunches) make up just 12 percent of high achievers but 36 percent of low achievers.

One key advantage of the OLDA data is that one can identify students who are high (or low) achievers early on and see what happens to them as they progress through school. This is another important piece of information that will be key to later analyses. To examine this, I convert the scores to SD units which show how a given student performs relative to his or her peers in the same grade and year. If we consider the average achievement of students in each third-grade quintile, those in the top quintile in third grade have an average math score of 1.34 SDs above the average student in that cohort while for the bottom quintile the score is -1.09 SD. However, by eighth grade, this gap compresses—for early high achievers, the average score drops to 0.90 SD and for low

[1] Although the data include students in third grade up through 2016–17, I restrict to these earlier cohorts so that they can be tracked through eighth grade. Note that through the rest of this report, I will refer to the academic year using the year of the spring semester (e.g., 2005–06 is 2006).

[2] For other outcomes not directly tied to math or reading, I consider a student to be a high achiever if he or she is in the top 20 percent of either the math or reading exams in third grade.

achievers it rises to -0.87 SD. ELA/Reading shows similar patterns, though the compression is not as great. We look more closely at these trends below.[3]

Finding 2: Only about half of students who score in the top quintile in third grade stay there, and the percentage is similar for the bottom quintile. Students in the middle quintiles stay roughly where they are.

Another way to look at these data is provided in Table 1, which provides a breakdown of the sample by third- and eighth-grade quintiles. For example, the cell in the far upper left of the top panel indicates that 11.3 percent of the sample are in the bottom achievement quintile in both grades three and eight. The cell in the far upper right indicates that 0.3 percent of the sample are in the bottom quintile in third grade and then move to the top quintile by eighth grade. If all students stayed at the same achievement levels relative to their peers, we would expect the diagonal elements (from upper left to lower right) of each table to equal 20 percent and all the off-diagonal elements to equal 0 percent. But that is not what happens: only about half of students who score in the top quintile in third grade stay there, and the bottom quintile is similar. In other words, there is achievement mobility in both directions. This is similar to what has been shown nationally (Xiang et al. 2011).

Table 1: Achievement quintile transitions between grades 3 and 8 (top panel, math scores; bottom panel, ELA scores)

		Eighth-grade quintile (math)				
		1	**2**	**3**	**4**	**5**
Third-grade quintile (math)	**1**	11.3%	5.9%	2.4%	0.9%	0.3%
	2	5.0%	6.9%	5.7%	3.3%	1.3%
	3	2.2%	4.5%	6.0%	5.4%	3.6%
	4	0.9%	2.4%	4.4%	6.0%	6.4%
	5	0.2%	0.7%	1.8%	4.0%	8.9%

		Eighth-grade quintile (ELA)				
		1	**2**	**3**	**4**	**5**
Third-grade quintile (ELA)	**1**	10.8%	5.1%	2.0%	0.7%	0.2%
	2	5.7%	6.6%	4.7%	2.5%	0.8%
	3	2.5%	5.0%	5.7%	4.9%	2.4%
	4	1.1%	2.9%	4.7%	6.2%	5.6%
	5	0.4%	1.2%	2.8%	5.6%	9.8%

It is important to understand that some student-achievement changes are due to random reasons from year to year. Because those in the top quintile cannot move higher, when measured this way, any students who fall in the achievement distribution are not offset by those who improve. With that in mind, it may be more instructive to look at what happens to students in the middle of the distribution—e.g., those in the third quintile in third grade. In this case, the results are rather symmetric; roughly similar shares of students improve and worsen, relative to their peers. Note, too, that even if high achievers are falling in the distribution, this does not necessarily mean that their learning suffers, as these are relative measures. For example, a student can remain in the same percentile—say, the fifty-fifth percentile—in third and eighth grade, but his or her scale score (which

[3] These results can be found in Panel B of Table A1, along with additional means and SDs of various measures across third-grade achievement quintiles in Panels C and D.

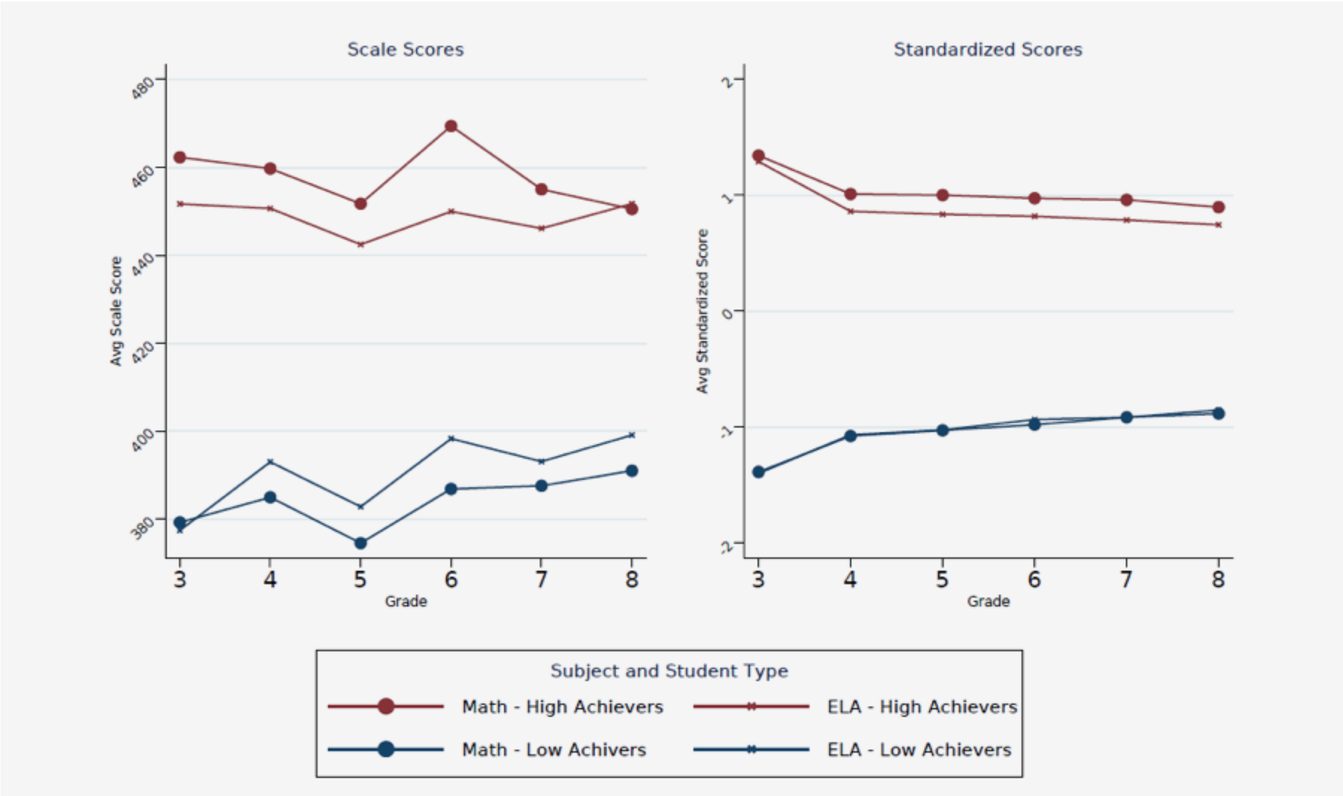
reflects whether a student is above or below grade level and hence is an absolute measure of achievement) could increase by fifty points if the average increase in scale scores in the cohort is fifty points.

Achievement trends through eighth grade

Next, we examine how achievement among early high achievers changes over the course of their educational careers. This study focuses on high achievers in particular, so in addition to investigating the extent to which they maintain their status through high school, we examine the how much this differs by demographic characteristics. Later, we will turn to studying how GT programs can serve to potentially help these students stay on top.

Figure 1 begins by showing overall trends in achievement for high versus low achievers based on third-grade achievement as students progress through school. On the left, I show the scale score of the OAA exams in grades 3–8. These scores measure achievement relative to an absolute standard. A student is considered proficient if he or she scores at 400. Here we see that low achievers show clear improvement over time, rising from a score of around 380 in both math and reading/ELA to around 410 in math and 400 in reading/ELA by eighth grade. For high achievers, however, scale scores remain relatively flat. On the right-hand side, the scale scores are converted to standardized units, which will be used throughout the rest of this report. A drawback of this measure is that by defining high and low achievers based on third-grade achievement, changes may simply reflect a statistical pattern called mean reversion, where temporarily high or low achievement tends to come back to the long-run average.[4] This pattern is apparent when looking at the change between grades 3 and 4 as the scores for high achievers drop while those for low achievers rise. Afterward, both trends level out. For this reason, I will focus on the trends from grade 4 through 8, as the mean reversion appears to taper out by grade 4.

Figure 1: Achievement over grade level, by high and low third-grade achievement



Note: Figure shows achievement by bottom and top quintiles of end-of-year third-grade exams and grade level.

[4] See, for example, Appendix I in Xiang et al. (2011).

Finding 3: Between fourth and eighth grade, early high achievers see their performance fall slightly or remain flat while the performance of low achievers increases.

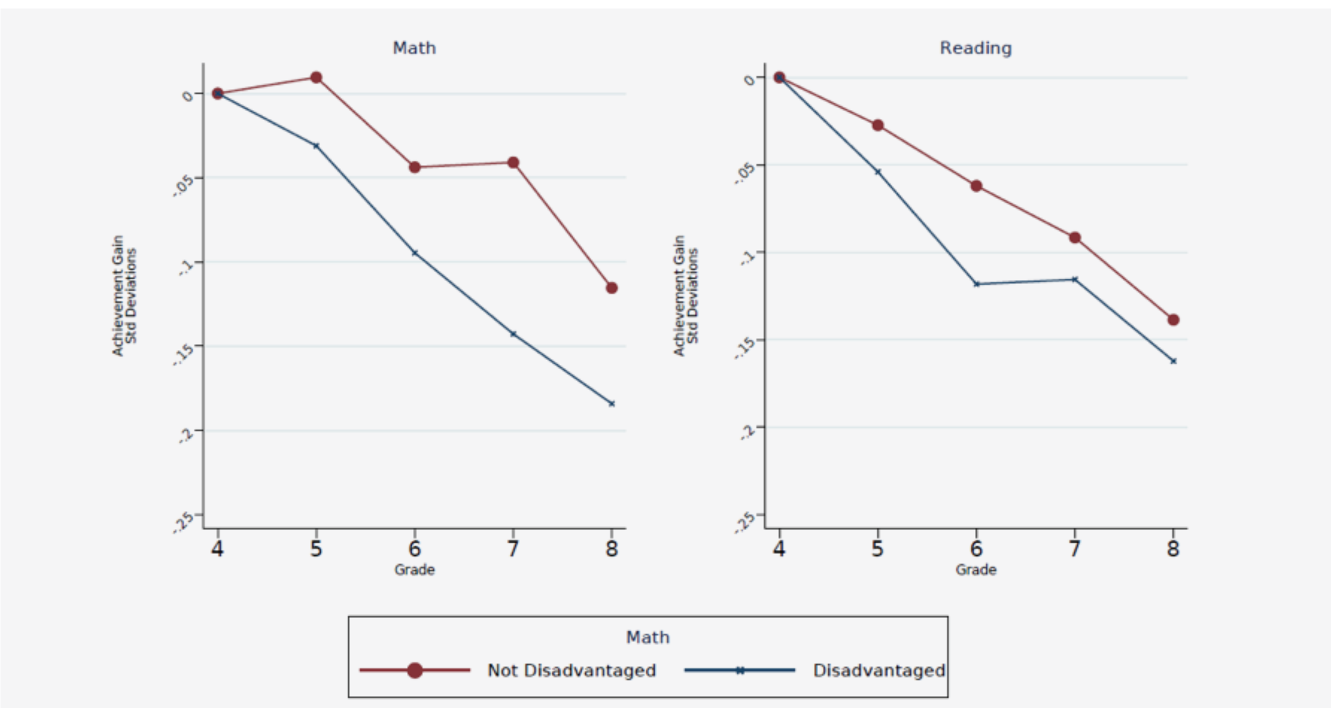
From Figure 1, we see that math scores for early low achievers increase between grades 4 and 9 by 0.06 SDs per year, while scores for early high achievers fall by 0.02 SD per year. Over the course of the four-year period, the increase for low achievers is sizable—approximately the same as replacing an average teacher (as measured by value added) with a ninety-fifth-percentile instructor for one year (Chetty, Friedman, and Rockoff 2014).[5] Comparatively, the drop in math achievement for high achievers is quite small. For ELA, the difference between the two groups is more substantial, with a 0.08 SD annual decrease for high and a 0.05 SD annual increase for low achievers, respectively.

Finding 4: Achievement growth among high achievers is lower for economically disadvantaged, female, Black, and Hispanic students. Growth is higher for API students.

What about achievement gains by various types of students? In Figures 2 through 4, I plot images of cumulative achievement gains for high achievers, separated by students' economic disadvantage status, gender, or race. The figures start at grade 4 and focus on gains relative to each student's fourth-grade achievement. Hence, for example, a value of 0.05 in seventh grade means that, on average, student achievement for that group increased by 0.05 SD between grades 4 and 7.

Figure 2 shows these results by economic-disadvantage status. We see that in both math and reading, disadvantaged students lose ground to nondisadvantaged students, though the gap is wider in math. Between fourth and eighth grade, their math achievement drops by 0.18 SD, while for nondisadvantaged high achievers, it falls by only 0.11 SD. Hence, the disadvantaged gap for high achievers widens as students progress through school.

Figure 2: Cumulative achievement gains for high achievers relative to fourth grade, by economic disadvantage

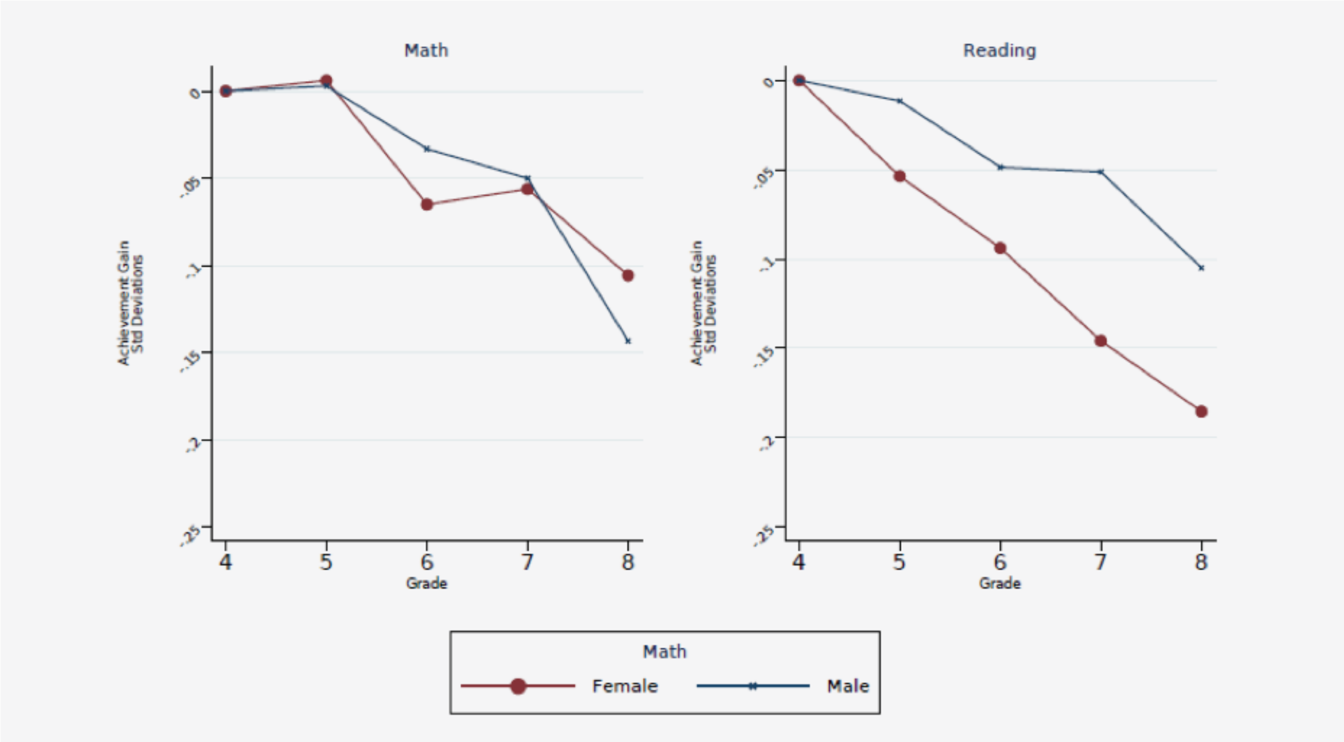


Note: Figures 2–4 show the cumulative achievement difference relative to the year after first third grade year (fourth grade for simplicity) for a student in the top quintile of third-grade end-of-year exams.

5 Calculation uses the estimate in Chetty, Friedman, and Rockoff (2014) that a one SD increase in teacher value-added increase student achievement in math by 0.14 SDs.

In Figure 3, we show the same trends by gender. Notably, there is little difference in achievement patterns for math between early-high-achieving girls and boys. This is a bit surprising given, existing evidence in other contexts that males tend to outperform females in math when restricted to high performers (Ellison and Swanson 2010). A relatively wide gap emerges in reading, where achievement growth for boys outperforms girls substantially. By the time they reach eighth grade, girls' performance in reading falls 0.07 SD more than boys' performance.

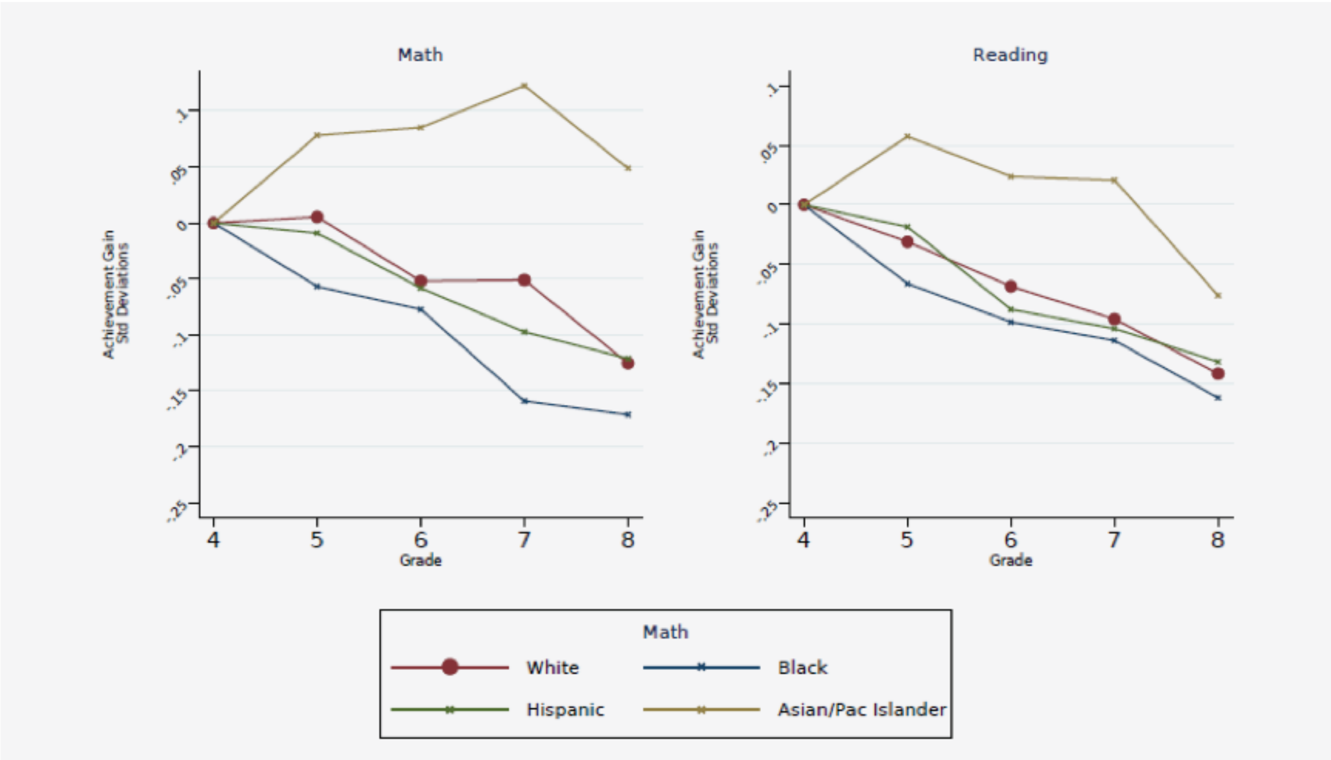
Figure 3: Cumulative achievement gains for high achievers relative to fourth grade, by gender



Finally, in Figure 4, we show achievement trends by race. Here, we see starker differentiation. API students do best at maintaining their performance relative to other races. In fact, they supply the only case where achievement (math) rises above the fourth-grade level.[6] Reading increases through seventh grade and then falls off dramatically. For other racial categories, achievement growth is similar for White and Hispanic students, but Black students lag slightly in both subjects through eighth grade. Overall, by the time these students reach high school, I find little difference in cumulative gains except for API students in math. One key aspect of all these figures, though, is that they show high-achieving students generally experience negative achievement growth relative to lower achievement peers, such that by the end of eighth grade, they lose around 0.1 SD in math and 0.15 SD in reading/ELA relative to their fourth-grade scores.

[6] When we measure relative to third grade, the math growth is also positive, and so API results even overcome the mean reversion effects.

Figure 4: Cumulative achievement gains for high achievers relative to fourth grade, by race



Long-term outcomes for high achievers—high school and beyond

Research question: How do high school and college outcomes for third-grade high achievers in Ohio differ by gender, race, and economic status?

Next, the analysis turns to longer-term outcomes among high achievers. The cohorts that can be analyzed are more limited and restricted to the 2006–08 third-grade cohorts for the examination of high school outcomes. This ensures that if students followed standard grade progression, the analysis observes through twelfth grade. The data provide information on test taking and scores for the ACT and the SAT, the two dominant college-entry examinations in the U.S., and AP exams. As expected, in virtually all cases, the outcomes improve as third-grade achievement quintiles rise.

For college outcomes, I am only able to consider the 2006 third-grade cohort, as this is the only cohort with both math and reading scores plus at least two years of post–high school observations. The college data come from the National Student Clearinghouse (NSC) and are matched to students by CHRR at Ohio State University. These include enrollment in two-year colleges (primarily community college and for-profit institutions) and four-year colleges, broken down by whether the four-year institution is public or private. Again, as with the high school outcomes, each enrollment category increases as third-grade achievement quintiles rise.

High school achievement

Finding 5: During high school, disadvantaged and racial gaps among Ohio high achievers remain for standardized, college-entrance, and AP exams. For race, Black students are the lowest performers, followed by Hispanic, White, and API students, in that order. There are no sizable gender gaps along these metrics, though girls are more likely to take SAT or ACT exams.

Figure 5 starts by showing differences in performance on Ohio's former high school exit exam, the Ohio Graduation Test (OGT), in math and reading. The figure shows the distribution for each group via box plots that show the mean (solid horizontal line in the center) and the interquartile range (the shaded boxes).[7] We see that, among high achievers, there is a slight gap between disadvantaged and nondisadvantaged[8] students, while there is virtually no difference in performance between males and females. However, there are large racial gaps. In both math and reading, Black early high achievers fare worst, followed by Hispanic students, White students, and finally API students, like what was found for achievement in elementary and middle school.

[7] Numerical values for these and other figures are available by request.
[8] Because disadvantaged status can change annually, I use whether the student was disadvantaged at any point during high school.

Figure 5: OGT standardized scores for third-grade high achievers (top panel, math; bottom panel, ELA)

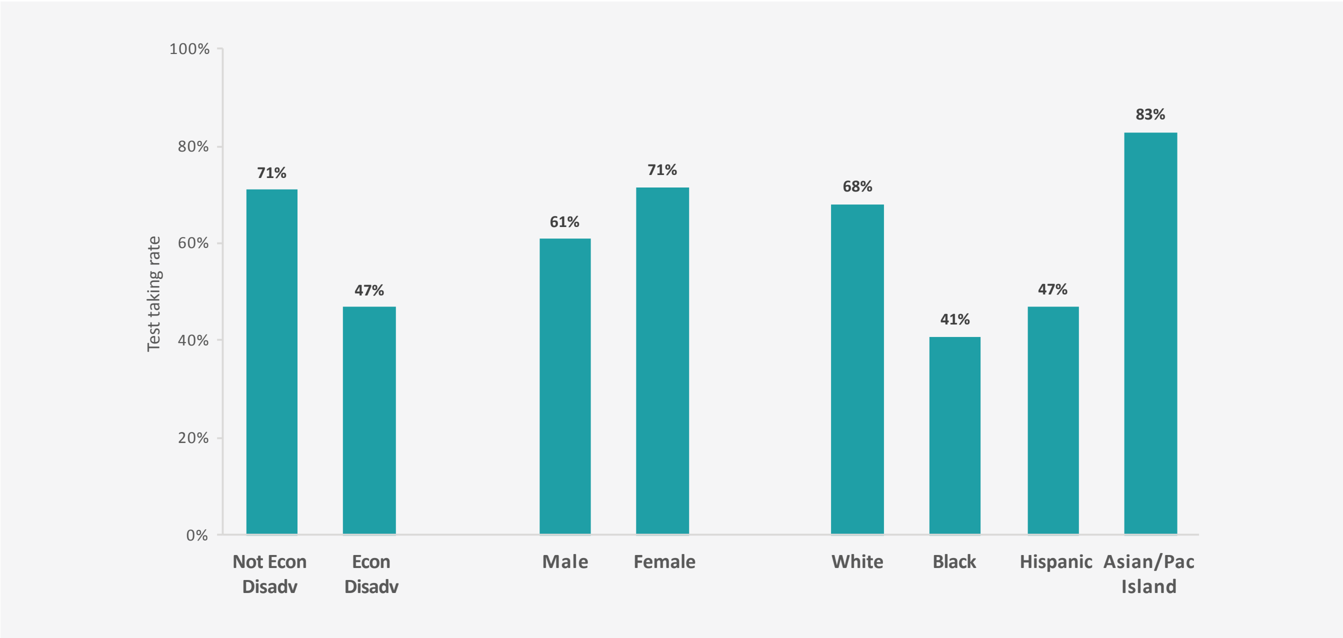
Note: Sample includes all students in 2006–08 third-grade cohorts who scored in the top quintile of the third-grade exam in the same subject and are observed in an Ohio high school. Figures show the mean value and interquartile ranges.

SAT, ACT, and AP Exams

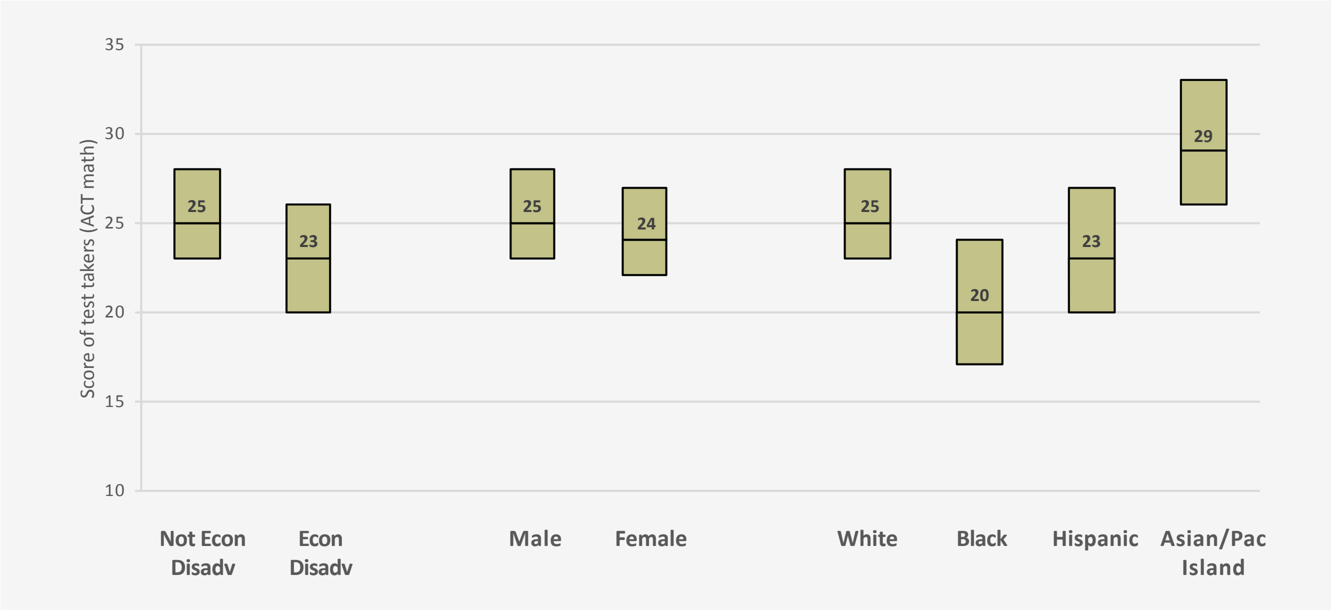
What about other outcomes that are particularly important to this subset of students, as they are precursors for college enrollment—i.e., SAT/ACT and AP exams?[9] In Figure 6, I show mean and interquartile ranges of outcomes on the ACT math and reading exams for students who were high achievers in third grade in the same subject, along with the average test-taking rates by demographic characteristics. The numbers for SAT exams are similar and provided in Appendix A (Figure A2). The patterns are similar to those seen in the OGT results. Disadvantaged students fare poorly relative to high-achieving nondisadvantaged pupils, and there are large racial gaps, with Black students showing the lowest performance and rates of test taking and API students the highest. Note that test taking is optional, and hence, scores are only available for those who choose to take the test. This likely makes these underestimates of the true gaps, as test-taking rates are lower among disadvantaged and Black or Hispanic students and these marginal nontakers would likely score lower than the average for their demographic.

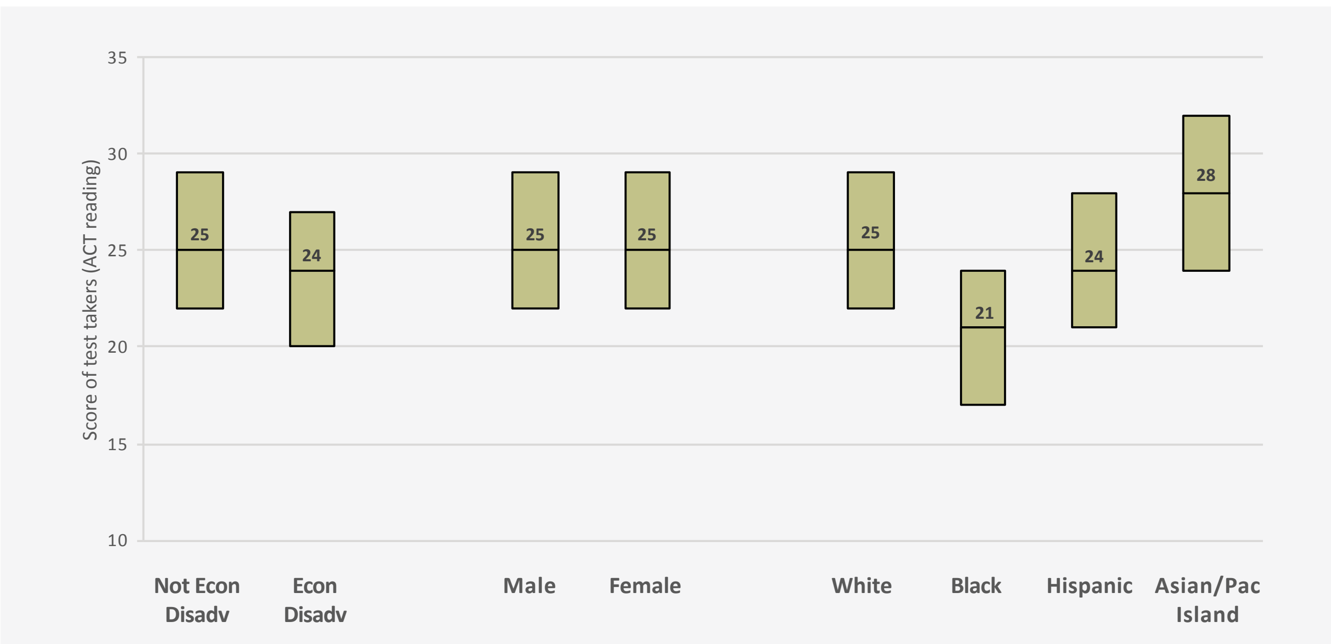
The main difference from Figure 5 is that a gender gap does appear to open in math, with male students performing better than female students. Recall that this is despite all these students being in the top quintile of third-grade achievement in that subject. Nonetheless, this could be due to the fact that, as we also see in the figure, girls are about ten percentage points more likely to take the exam. If these women who would not have taken the test if they were men tend to be lower performing than the average male, this would make a gender gap appear when it is just due to these women providing scores that would not be observed if they were men.

Figure 6: ACT scores for third-grade high achievers (top panel, test-participation rates; middle panel, ACT math; bottom panel, ACT reading)



9 The ACT exam is more commonly used in Ohio than the SAT exam, but some students take both tests.





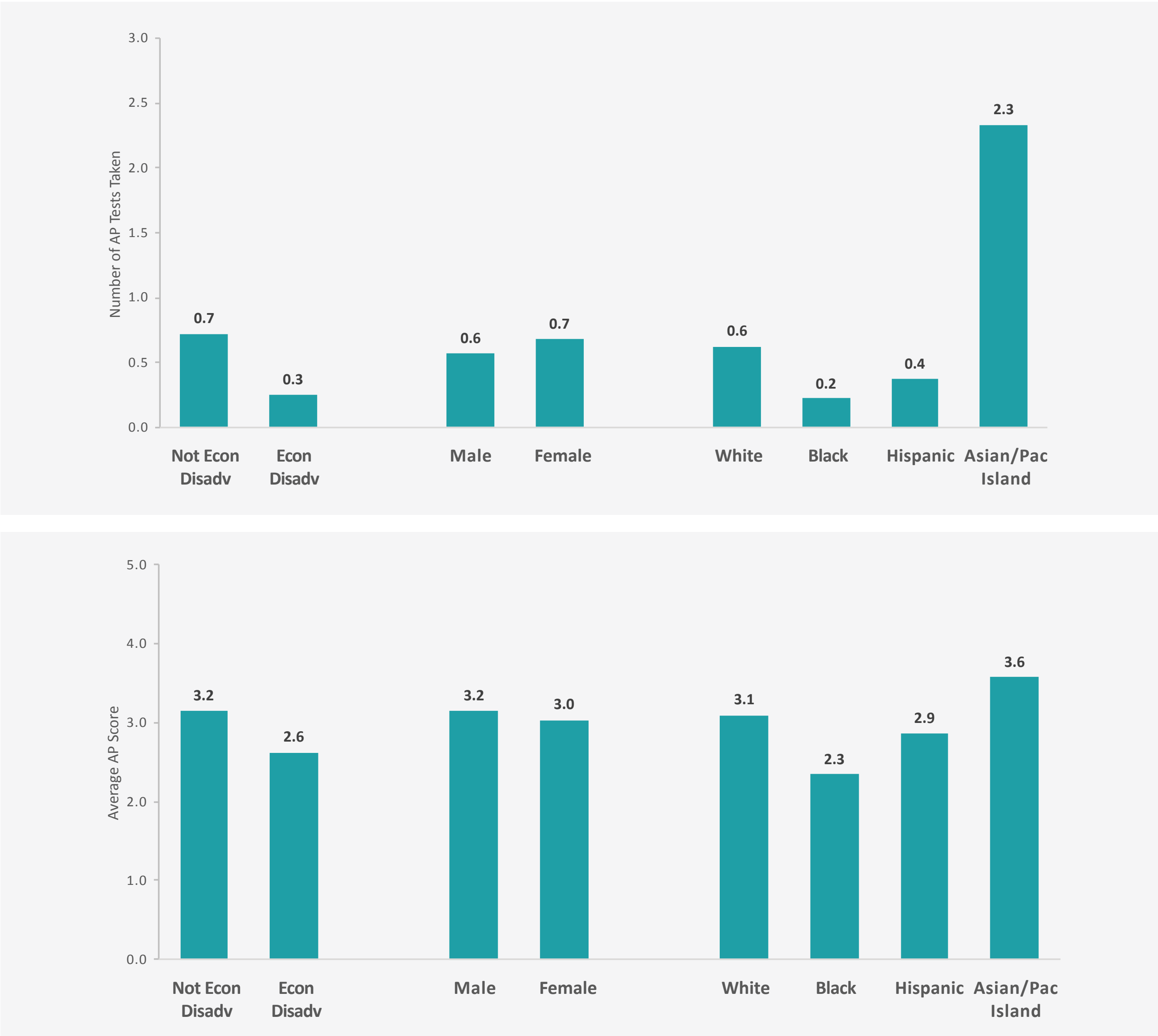
Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of third-grade exams in same subject (math for test taking) and are observed in an Ohio high school. Figure shows mean values and the interquartile ranges. SAT exam data are provided in the appendix.

Another useful measure of long-run performance of high achievers is participation and performance in the AP program.[10] Data for students in Ohio are recorded on up to five exams. Figure 7 shows the average number of tests taken and the average score for exams taken by students who were high achieving in third-grade math by demographic category.[11] The patterns for average scores mirror those for the OGT and SAT/ACT exams. For the number of tests taken, the direction of the gaps remains the same—economically disadvantaged, male, and Black and Hispanic students take fewer exams—but API students take many more exams than others. Such deficits in participation in AP programs could hamper future success, as research shows that AP programs increase college matriculation (Jackson 2010).

[10] A similar program commonly used in some schools is the International Baccalaureate Program. Unfortunately, I do not have data for this program.

[11] Means for high achievers in reading are very similar and provided in Figure A4.

Figure 7: AP exams for high achievers in third-grade math (top panel, average number of AP exams taken; bottom panel, average AP exam score)



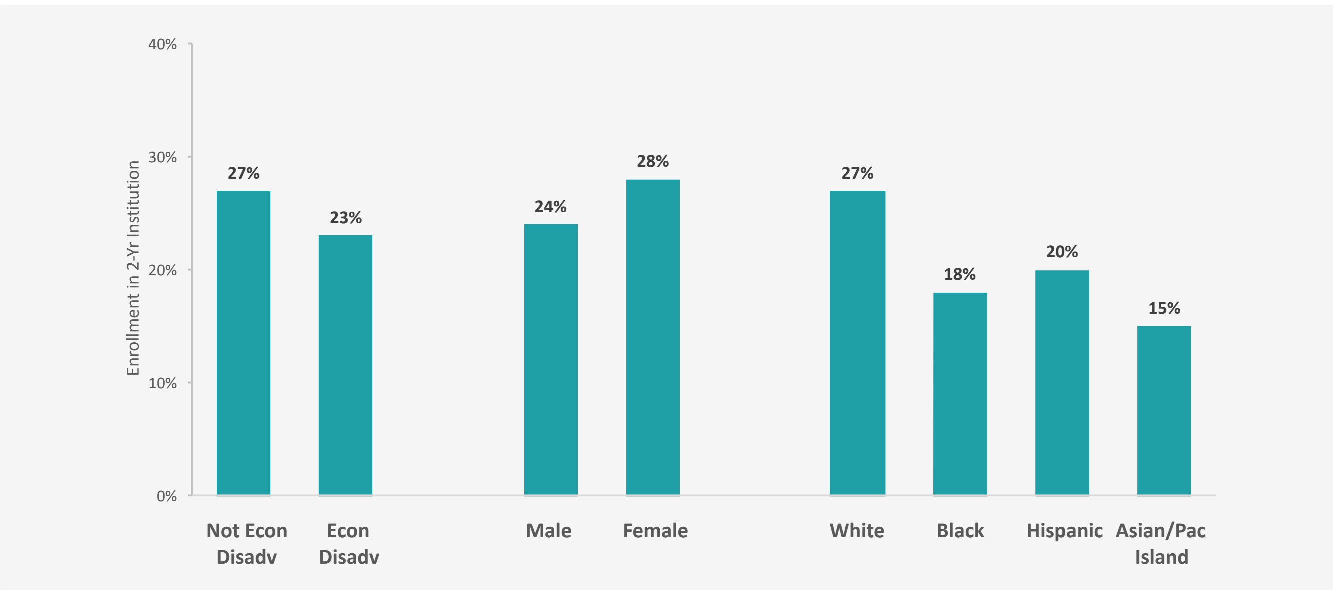
Note: Data are restricted to the 2006–08 third-grade cohorts.

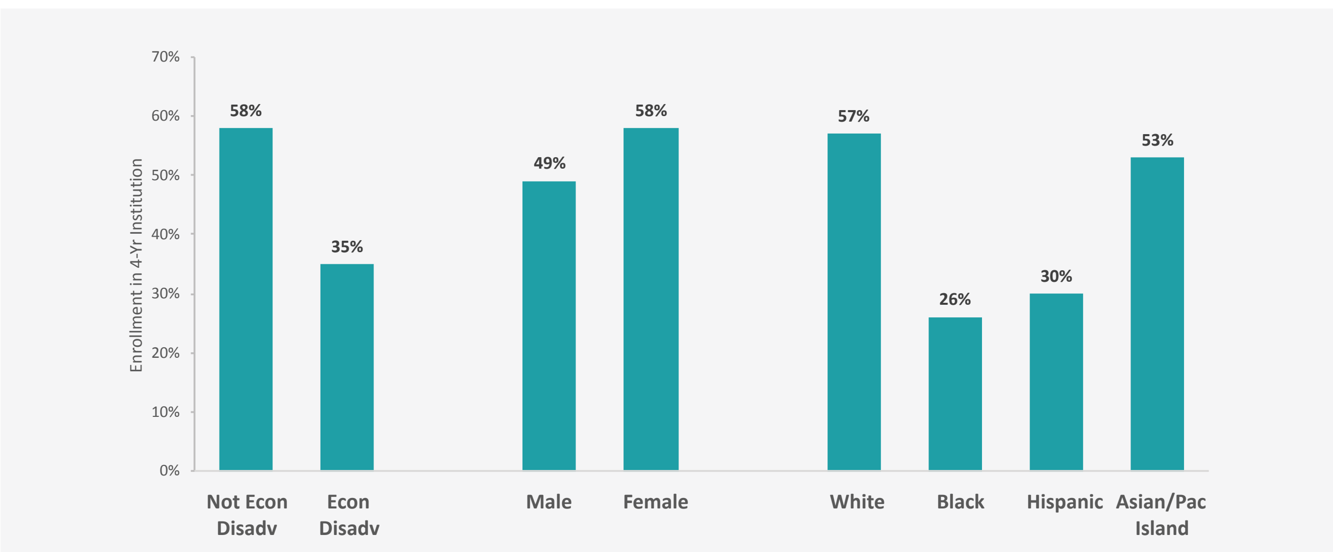
College enrollment

Finding 6: Similar gaps by race and income among early high achievers emerge when considering college enrollment, with the notable exception that enrollment in two-year schools is lowest for API high achievers; while they have higher enrollment in four-year schools, it nonetheless slightly lags that for White students. Further, female enrollment in both types of colleges exceed male enrollment.

After looking at high school outcomes, I turn to college outcomes and see the same general patterns emerge. The results are less precise because only one cohort can be tracked from third grade through two years after high school graduation. Figure 8 shows college-enrollment rates for high math achievers split by two-year and four-year schools. Results for early high achievers in reading are provided in the appendix (Figure A6) and show similar patterns. Enrollment rates in two-year schools are lower for disadvantaged students and males. Across racial categories, the highest enrollment rates are for White students, followed by Black, Hispanic, and finally API students. For four-year institutions, we see a similar pattern, but the differences are larger. The exception is for API students, who are only slightly behind White students in four-year schools. The difference in four-year enrollment between White and Black or Hispanic students is particularly striking. Recall that these are all early high achievers, so they scored in the top 20 percent in math in third grade. Even so, there are massive differences in college enrollment. Of the White early high achievers, 57 percent enroll in a four-year school within two years after normal grade progression would imply that they graduated high school. However, the rates for Black and Hispanic early high achievers are around half that, at 26 and 30 percent, respectively. The gaps by economic status are similarly large, with 58 percent of nondisadvantaged students attending four-year colleges but only 35 percent of disadvantaged students enrolling. The foregoing analysis does not consider the quality of a college or university; that issue, however, is explored in Appendix B.

Figure 8: College enrollment for high achievers in third-grade math (top panel, enrollment in two-year schools; bottom panel, enrollment in four-year schools)





Note: Data are restricted to the 2006 third-grade cohort. Enrollment is based on any institution attended within two years of expected high school graduation (eleven years after third grade).

Although the patterns for income and gender persist after we account for third-grade achievement and within-school factors, accounting for these changes the story for race considerably. As shown in the appendix (Figure A7) and discussed in the sidebar below, Black and Hispanic enrollment rates are slightly *higher* than for White students in both two-year and four-year schools, but these differences are small and the only one that's significant is Hispanic students for two-year colleges (five percentage points). Meanwhile, API students show statistically significantly *lower* enrollment than White students in both types of institutions (nine percentage points in two-year and six percentage points in four-year colleges). These results suggest that college-enrollment gaps across races are mostly a function of differences across early schooling locations, rather than via students attending the same schools and living in the same locations. This is particularly interesting considering recent evidence that geography is a key factor in long-term success (Chetty and Hendren 2018a, b; Chyn 2018; Chetty, Hendren, and Katz 2016; Chetty et al. 2014); that is, where you live matters quite a lot in terms of long-run outcomes. The results here are consistent with those findings. That said, it is important to emphasize that where one grows up—and associated correlates such as differences in school resources that could contribute to lower readiness for college—is unlikely to be the sole factor at play, and there are many important differences in terms of the educational environments and opportunities provided to high achievers from different demographic groups that likely do play important roles.

Regression analyses

Finding 7: Accounting for differences across high-achieving students in their precise third-grade achievement levels, elementary school attended, and cohorts produces few changes in the gaps seen in the underlying unadjusted data, except for college attendance, where API students fare worse and other races have similar enrollment rates.

To provide a deeper analysis of these differences by demographic groups, I provide results from a series of regression models in the appendix—Figures A1, A3, A5, and A7 (starting at page 52).[12] The figures show estimates from a regression of the outcome on indicators for the different demographic groups. Hence, the estimate for female students is the difference in achievement relative to male students; the estimates for Black, Hispanic, and API students are relative to White students; and the estimate for economically disadvantaged students is relative to nondisadvantaged students. Further, as noted in the Methods section above, I control for some important factors that ensure that each student is compared to another student who had a similar educational environment in third grade and started with the same test scores. This way, the analysis allows us to make apples-to-apples comparisons of students who look similar on observed dimensions, at least when they were in third grade.

Figure A1 shows these demographic differences for high school test scores. This figure, and the figures that follow, show estimates from linear regressions as described above where the "dot" shows the estimate itself and the "whiskers" show the 95 percent confidence interval, which roughly translates into there being a 95 percent probability that the true value lies somewhere between the whiskers. Hence, if the whiskers are both below or both above a value of zero, that means we can say that the specified category has statistically significantly different achievement from the left-out category. Hence, the estimate for economically disadvantaged students is interpreted as being relative to nondisadvantaged students, the estimate for female students is relative to male students, and the estimates for Black, Hispanic, and API students are all relative to White students.

On all five subjects—math, ELA, science, social studies, and writing—API early high achievers have the highest achievement, followed by White students, then Black students, and finally Hispanic students. Economically disadvantaged high achievers perform worse than nondisadvantaged in all subjects, as well. For gender, on the other hand, males perform better than females in math, science, and social studies, but the reverse is true for writing.

[12] Technical details are provided in Appendix A.

Figure A3 provides similar estimates with SAT and ACT scores as the outcomes. Disadvantaged students are less likely to take the exams and, among those who do take them, perform worse than nondisadvantaged students. Females are more likely to take an exam than males, but those who do perform worse on all except ACT Reading. However, this is likely partially due to the higher exam-taking rate drawing in more academically marginal women. The racial patterns largely match what we see in achievement; however, Black high achievers are more likely to take an exam than White high achievers, after accounting for precise third-grade achievement, elementary school, and cohort—a pattern that does not show up in the raw data in Figure 6.

In Figure A5, we see that AP exam taking is only slightly lower for disadvantaged relative to nondisadvantaged students, for males relative to females, and for Black students relative to White students. However, there is a large deficit in exam taking for Hispanic students and substantially higher exam taking for API students relative to White students. For the average scores on exams, the patterns are generally similar to what we see for SAT/ACT exams in Figure A3.

Finally, in Figure A7, I provide regression estimates for college enrollment. Disadvantaged high-achieving students attend two-year schools at similar rates to nondisadvantaged high-achieving students but lag far behind in four-year school enrollment, while male students lag female students in both cases. The racial differences, on the other hand, are somewhat surprising. After accounting for the factors mentioned above in the regressions, White, Black, and Hispanic students attend college at roughly similar rates (with Hispanic students having slightly higher two-year enrollment), but API students lag far behind their high-achieving counterparts in college enrollment, despite their higher performance on college-entrance and AP exams. This highlights a potential hidden problem, where high-achieving API students are not enrolling in college at similar rates to their high-achieving peers who grew up in the same locales.

In sum, there is substantial variation in long-run outcomes among demographic groups for early high achievers in Ohio across demographic groups. This generally holds whether we look at raw data or compare people who attend the same elementary school cohort and have the same third-grade test scores. Though magnitudes differ by outcome measure, early high achievers who are economically disadvantaged generally have worse high school and college outcomes than nondisadvantaged students. For racial categories, Hispanic high achievers in Ohio appear to fare worst in their long-run outcomes, followed by Black students, White students, and finally API students, who tend to perform the best. This pattern is consistent across high school achievement, ACT and SAT exam outcomes, and AP exam outcomes. Finally, gender impacts are mixed and tend to vary by subject and outcome.

Gifted and talented programs in Ohio

We now turn to GT programs, which could play a role in improving the later performance of early high achievers and perhaps closing some of the gaps we saw above. These programs are commonly available for some, but not all, high achievers, and in Ohio (as elsewhere), the standards for enrollment in GT programs are restrictive. Hence, many high achievers, as defined in this report, are never identified as gifted. Conversely, some GT students are not early high achievers, as defined in this report. In Ohio, about 15 percent of total public school enrollment being identified as gifted in 2010.[13] Using the OLDA data, in 2017, this number for students in eighth grade was 20 percent.

Since 1999, Ohio law has required public noncharter school districts to identify students as gifted in any of four areas: (1) superior cognitive ability; (2) specific academic ability in math, science, ELA, and/or social studies; (3) creative thinking; and (4) visual or performing arts.[14] In 2017, 7 percent of eighth-grade students were identified via cognitive ability, 10 percent via math, 3 percent via science, 6 percent via reading, 3 percent via social studies, 1 percent via creative thinking, and less than 1 percent via visual or performing arts.[15] Students may be identified in multiple areas; for instance, a child may be identified as having both superior cognitive ability and a specific academic ability. This applies to about 27 percent of eighth-grade students with any identification in 2017.[16] Assessment for gifted identification may occur either through a referral process that is largely subjective, as it is typically initiated by teachers and/or parents[17] or via whole-grade screenings that occur at least once during grades K–2 and once during grades 3–6.[18] Gifted identification is permanent—a "once gifted, always gifted" policy—regardless of subsequent test or classroom performance.[19]

Although districts must identify gifted students, they are not required to provide services to all students who are identified, even though each district is required to develop a plan for gifted identification. In this plan, they may include several specified options set forth in state law, including a differentiated curriculum, cluster grouping, accelerated course work, and AP and international baccalaureate courses.[20]

13 Data on gifted identification over time and by district typology is available at https://www.ohiobythenumbers.com.
14 See http://codes.ohio.gov/orc/3324.03v1. Passed in 1999, House Bill 282 of the 123rd General Assembly enacted significant changes in gifted education law; see http://lsc.state.oh.us/analyses/fnla123.nsf/All%20Bills%20and%20Resolutions/527153332307BB07852567DD00456E1D.
15 Author's calculation from OLDA data.
16 Author's calculation from OLDA data.
17 See http://codes.ohio.gov/orc/3324.04v1.
18 See http://codes.ohio.gov/oac/3301-51-15. Universal whole-grade screenings were mandated statewide starting in 2017, but such screenings were performed at district discretion prior to that.
19 See http://codes.ohio.gov/oac/3301-51-15.
20 See http://codes.ohio.gov/orc/3324.07v1.

Demographic differences in gifted status for high achievers

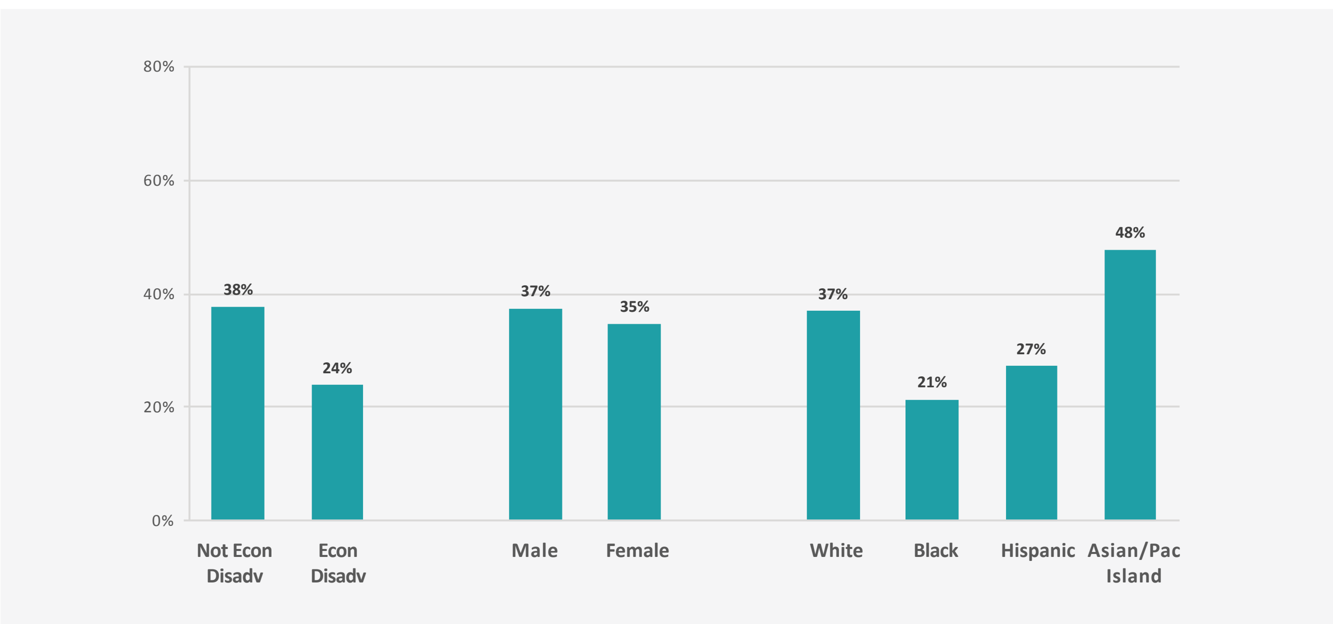
Research question: How does gifted status for high achievers differ by gender, race, and economic status, and how many high achievers are eventually identified as gifted?

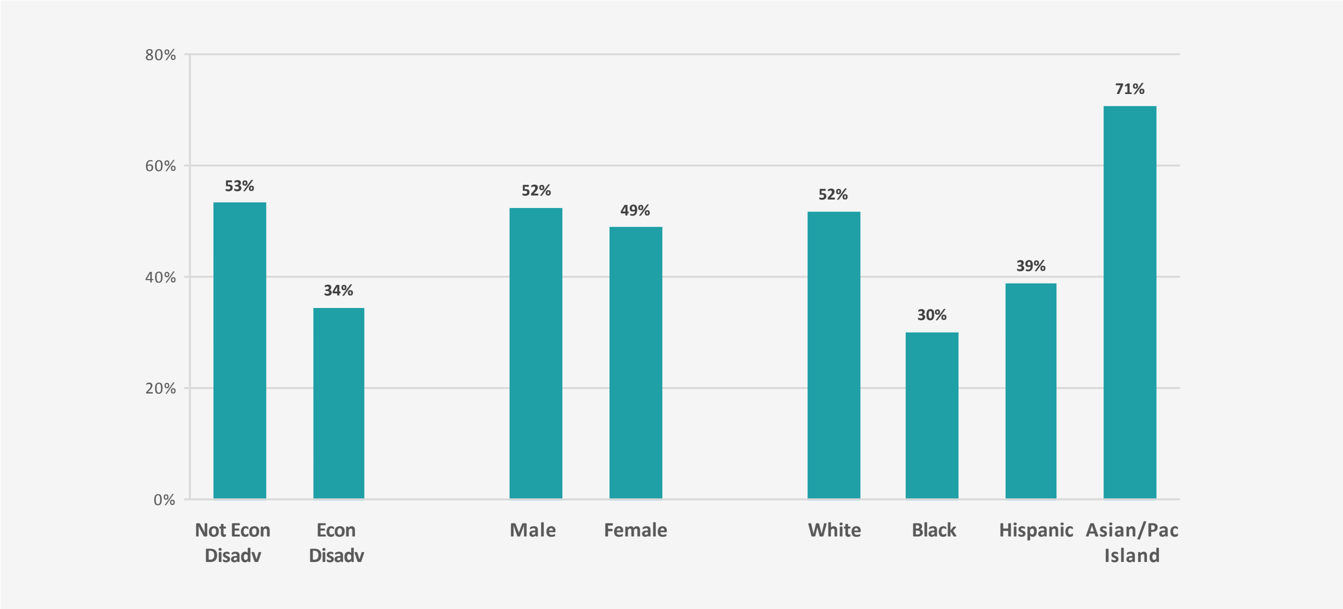
To look at the potential impact that GT identification could have on early-high-achiever performance, I again start by looking at some trends in achievement. These trends do not tell us anything about whether gifted identification has an academic impact—that will be addressed in the next section. They do, however, allow us to get a feel for the underlying framework within which gifted identification policies operate. It is also worth noting that in Ohio, during our period of study, service rates overall were generally low, between 20 and 30 percent of identified students during the period covered in this study. Since then, Ohio has drastically increased service rates to 60 percent of those identified as of 2019–20. Nonetheless, the low service rate during the period of analysis will be important context when trying to interpret estimates of the impact of GT on student outcomes.

Finding 8: High achievers who are disadvantaged, Black, or Hispanic are substantially less likely to be identified for gifted services.

To start, Figure 9 provides gifted rates in grade 3 and grade 8 for third-grade high achievers (defined by math or reading score) by demographic category. In general, GT identification rates between grades 3 and 8 increase by about fourteen percentage points, with relative gaps across groups remaining consistent over time. Note that this increase is largely expected, as once a student is identified for gifted services, he/she does not lose that identification. Nonetheless, the demographic gaps are quite large. First, nondisadvantaged high-achieving students are substantially more likely to be identified as gifted along some dimension by eighth grade (53 percent) than disadvantaged students (34 percent). Further, there are large racial gaps, with eighth-grade identification rates of 71 percent for API high achievers, 52 percent for White high achievers, 39 percent for Hispanic high achievers, and only 30 percent for Black high achievers. For gender, the values are roughly equal.

Figure 9: GT status by demographic and grade level for high achievers (top panel, third grade; bottom panel, eighth grade)





Note: Figure shows mean GT identification rates for each demographic group, given students score in the top quintile of their third-grade math exams.

Finding 9: Only 23 percent of high achievers in third-grade math in Ohio are identified for math GT services in third grade. By eighth grade, this only increases to 36 percent. In reading, 20 percent of third-grade high achievers are identified for reading GT services in third grade; this increases to 28 percent by eighth grade.

Next, returning to Table A1 in the appendix, I show in Panel D GT identification in math, reading, and any subject (other subjects include cognitive ability, science, creative thinking, social studies, and visual or performing arts) for third and eighth grade by third-grade achievement quintile. Not surprisingly, there is a clear positive relationship between achievement in third grade and GT identification. Nonetheless, the relationship is not perfect, indicating that achievement is not the only factor driving GT identification. Further, a large portion of high achievers are not identified as GT in the tested subject or any category. Only 23 percent of high achievers in math are identified as gifted in math in third grade, and this only increases to 36 percent by eighth grade. The numbers are smaller for reading, at 20 percent and 28 percent, respectively. Overall, just 57 percent of students who score in the top quintile in math in third grade are identified as gifted in at least one subject. At the same time, a small but substantial share of students with middling achievement in third grade are identified. Of students in the third quintile (fortieth to sixtieth percentiles) of third-grade math achievement, 9 percent are identified in third grade and 14 percent are eventually identified by eighth grade.

Is gifted identification associated with long-term outcomes?

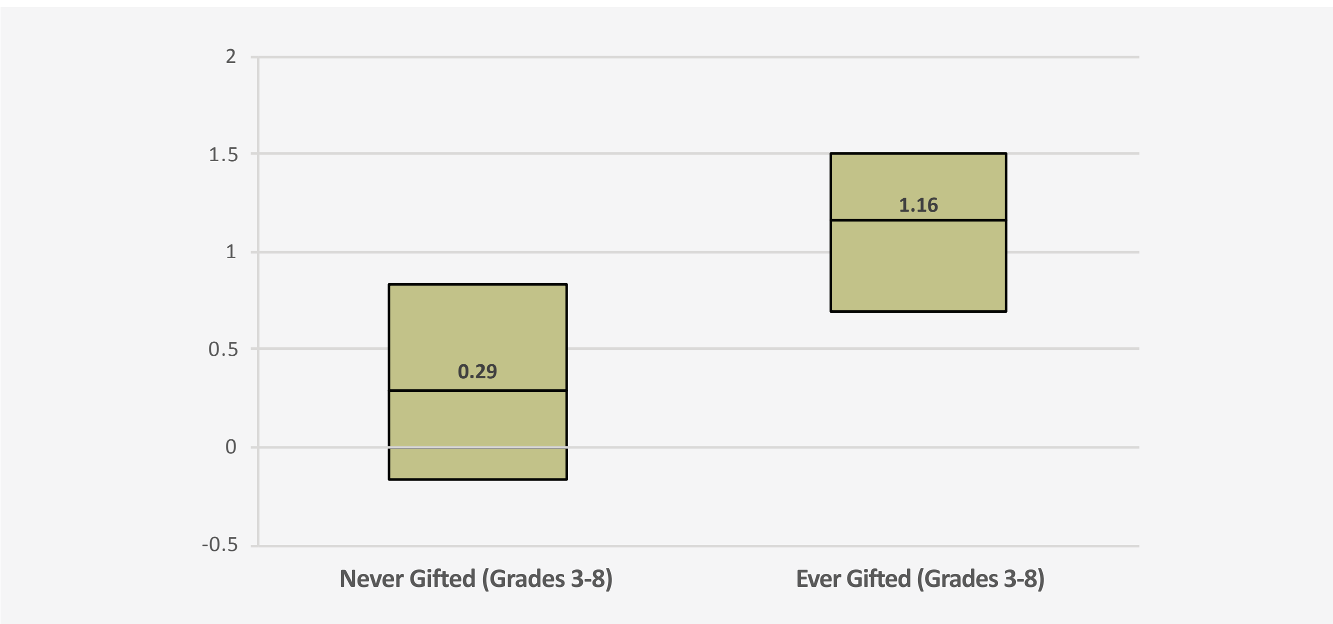
Research question: How do high school and college outcomes for third-grade high achievers in Ohio differ by gifted status?

<u>High school achievement</u>

Finding 10: Early high achievers who are identified for GT programs exhibit higher performance on high school achievement exams even when comparing those with similar early achievement and who are from the same elementary school, race, gender, and economic status—by 0.4 SDs in math and 0.3 SDs in ELA.

Starting with Figure 10, I look at how performance on the OGTs differs by whether a student was identified as GT between grades 3 and 8. As expected, early high achievers who are identified as gifted in elementary or middle school perform substantially better in high school. On average, early math achievers who are identified as GT in math score 0.57 SDs higher on the math OGT exam than early math achievers who do not receive such an identification. For reading/ELA, the gap is 0.39 SDs.

Figure 10: OGT standardized scores for third-grade high achievers by GT status (top panel, math; bottom panel, reading)



Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of the third-grade exam in the same subject and are observed in an Ohio high school. Figure shows mean values and interquartile ranges.

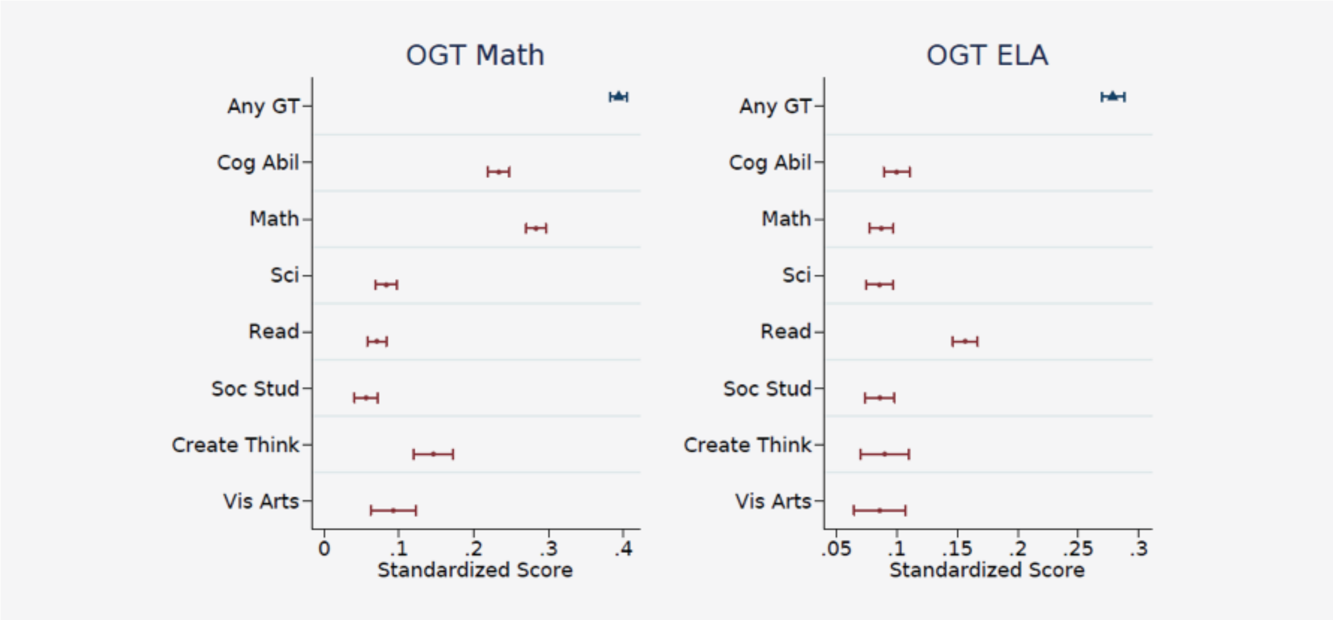
In Figure 11, I provide estimates for differences in GT identification using the same regression analysis described in the sidebar above but with additional variables for whether a student was identified as gifted at some point between grades 3 and 8 in each of the following categories: cognitive ability, math, science, reading, social studies, creative thinking, and visual or performing arts. To be considered a GT student, the student must be identified in one of these categories. Nonetheless, students often qualify for multiple categories. For high school students who were high achievers in third grade, 43 percent were not identified in any category between grades 3 and 8, while 17 percent were identified in only one category, 13 percent in two categories, and 28 percent in more than two categories. Details on the specific estimation model—and the models that follow—are provided in Appendix B. The estimates shown in the figure show the difference in OGT scores for a high achiever (based on reading or math) with the given GT identification, relative to high achievers who attended the same elementary school, are the same race and same gender, and have the same economic disadvantage status.[21]

Figure 11 shows that, as expected, even when controlling for third-grade achievement, demographic characteristics, and being from the same elementary school, students with GT identifications perform better on all OGT exams, regardless of their specific identification. The first row shows the results from a regression that includes a single GT indicator for whether the student qualified for GT under any category. Note that in these models, I define a high achiever as someone who scores in the top quintile of math *or* reading in third grade, and so the estimates are not directly comparable to the distributions shown in Figure 10. Nonetheless, we still see that GT high achievers perform substantially better than non-GT high achievers—they score 0.4 SD and 0.3 SD higher in math and reading.[22] The rest of the estimates in each panel show the results of a second model that instead provides indicators for each specific GT qualification. Note that these estimates are considerably smaller than the "any-GT" estimates. Because many students qualify under multiple categories, each individual category will only partially contribute to the overall result. In general, the estimates for specific GT qualifications are between 0.05 and 0.23 SD and are always statistically significantly larger than zero.

21 Though it is tempting to consider these estimates to be additive—e.g., that the difference in achievement between a student who is identified as GT in math and reading and a student with no identification is the estimate on math plus the estimate on reading—this would not be an appropriate interpretation, as the model is not structured to pick up these combined effects.

22 In Figure A11, I provide results from models that look at science, social studies, and writing. The estimates are similar at 0.25 SDs in writing and around 0.4 SD in science and social studies.

Figure 11: Estimates of GT differences in high school achievement with demographic controls for third-grade high achievers



Note: Data are restricted to students in 2006–08 grade 3 cohorts who score in the top quintile of math or reading in third grade. Category of GT identification is listed on the y-axis. Students can be identified in multiple categories. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects. The estimate for "Any GT" is from a separate regression.

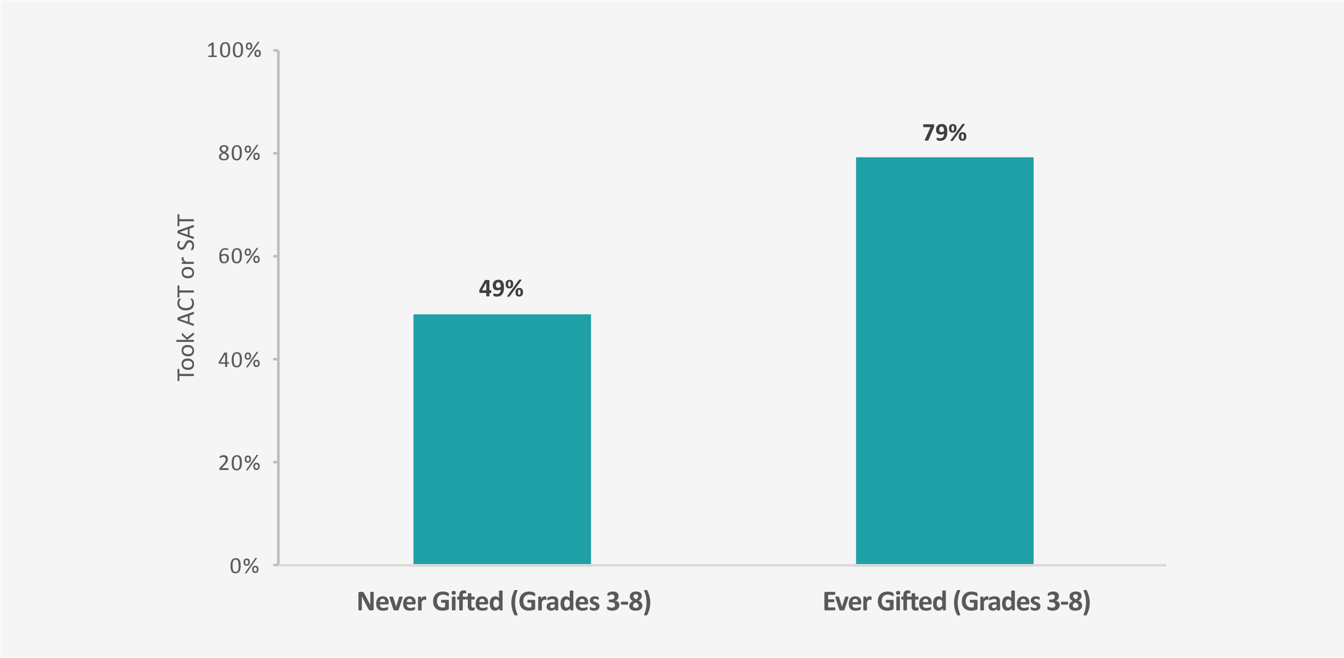
SAT, ACT, and AP Exams

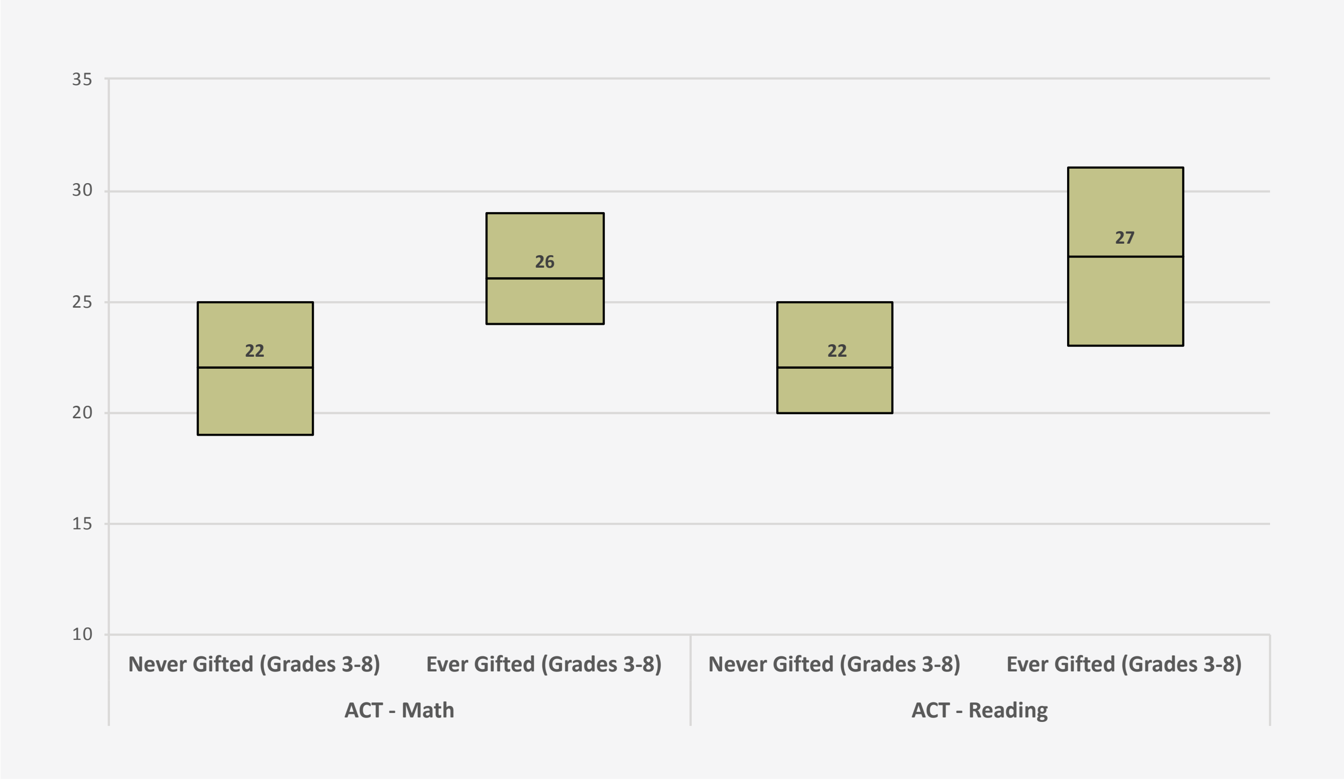
Finding 11: High achievers who are identified as GT are more likely to take a college-entrance exam and to score higher on it. Similarly, they are more likely to take AP exams and score higher. These hold after controlling for early achievement, being in the same elementary school, race, gender, and economic status.

Figure 12 provides distributional information about ACT exam taking and ACT performance by GT status for high achievers in the given subject.[23] As with the OGT scores, high achievers who are identified as GT are more likely to take a college-entrance exam and to score higher on it.

[23] Math is used to identify high achievers when looking at whether a student took any exam. Differences in SAT scores are similar and shown in Figure A12.

Figure 12: ACT scores for third-grade high achievers by gifted status (top panel, test participation rates; bottom panel, average ACT Math and ACT Reading scores)

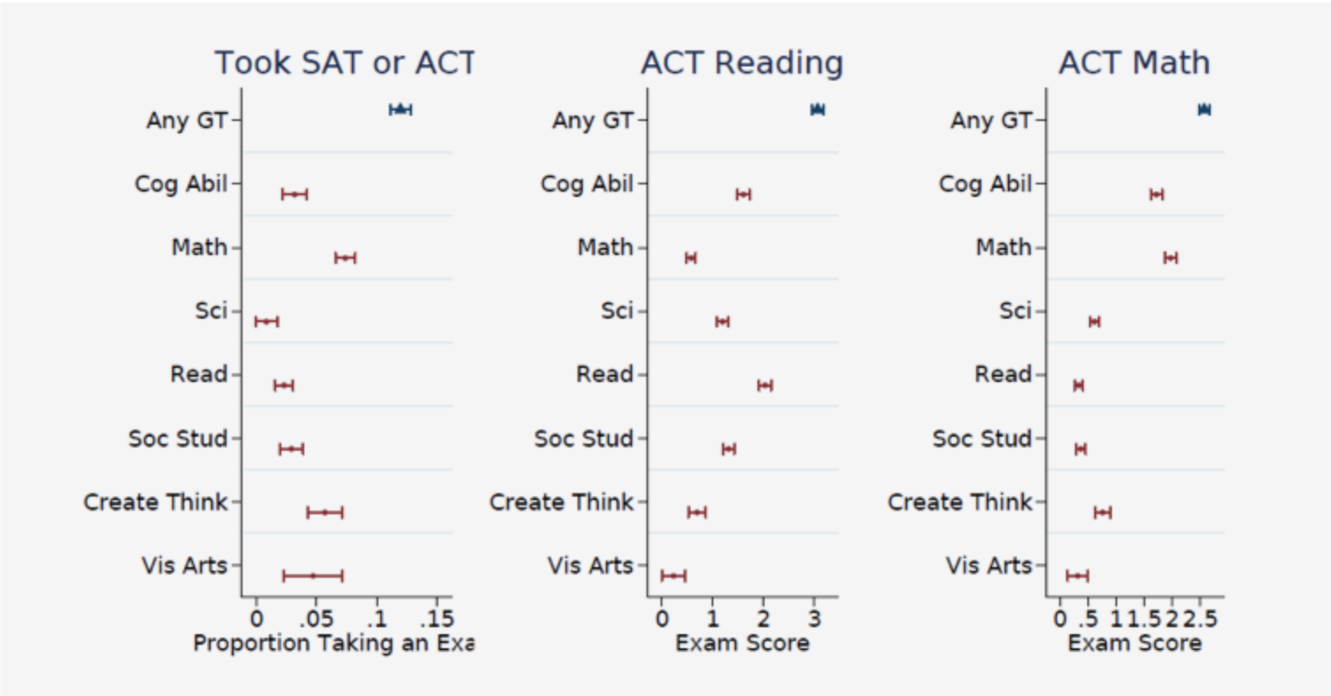




Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of third-grade exam in same subject (test taking uses math quintile) and are observed in an Ohio high school.

Figure 13 shows regression results that corroborate the raw differences. After accounting for third-grade achievement, demographics, and elementary school cohort, students identified as gifted remain more likely to take the SAT or ACT and score higher than other high achievers in most cases. Based on these estimates, students with any gifted identification are twelve percentage points more likely to take either exam and score three points higher (out of thirty-two) on the ACT and sixty points higher (out of 800) on the SAT.[24]

Figure 13: Estimates of GT differences in ACT exams with demographic controls for third-grade high achievers

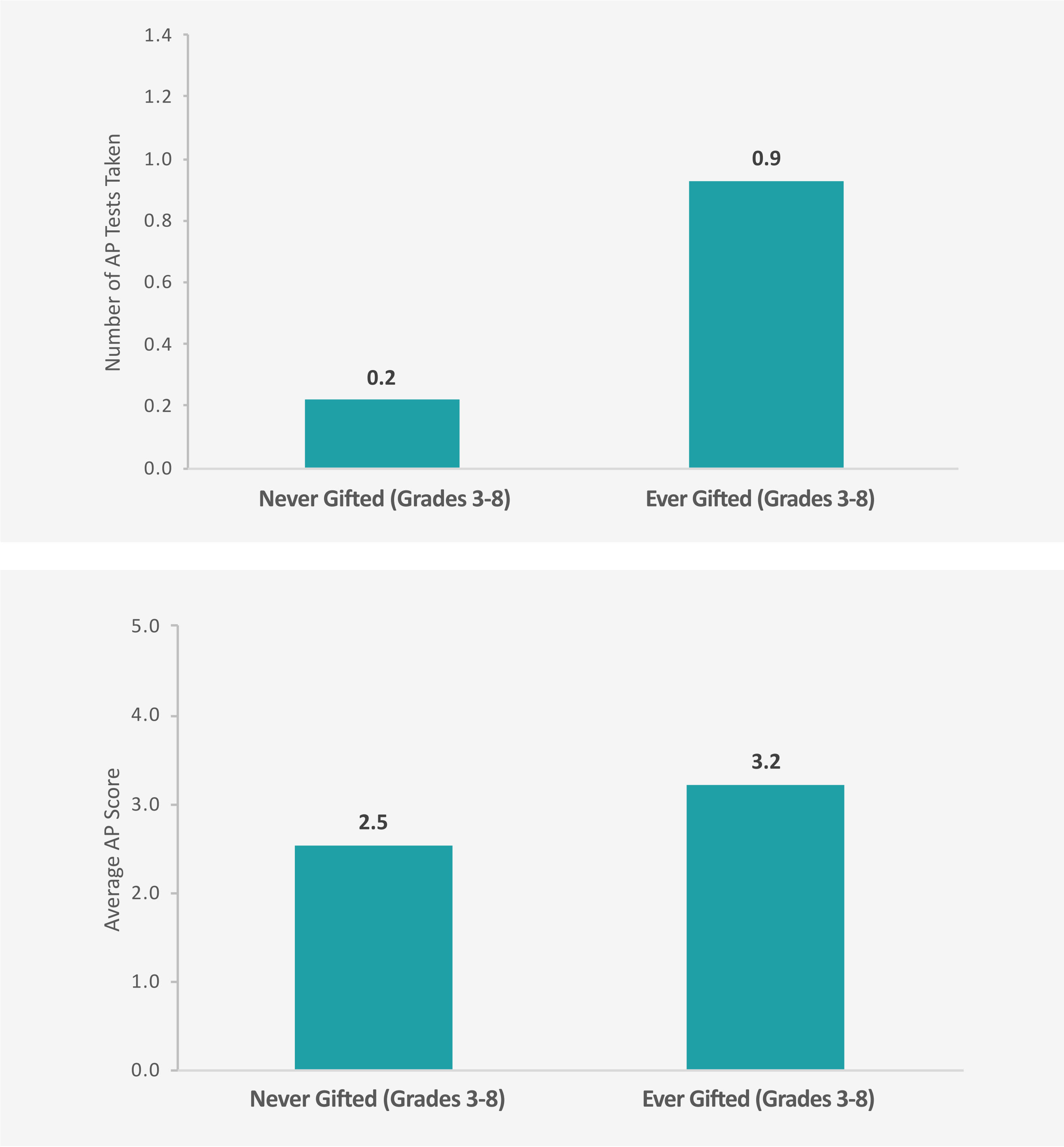


Note: Restricted to students in 2006–08 third-grade cohorts who score in the top quintile of a given subject ("proportion taking an exam" uses math) in third grade. Category of GT identification is listed on the y-axis. *Students can be identified in multiple categories. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects. The estimate for "any GT" is from a separate regression.*

In Figure 14, I investigate the relationships between gifted identification and AP-exam outcomes for early high achievers in math (those for early high achievers in reading are provided in Figure A14). GT students take more AP exams and score higher on the exams they take than other high achievers. Figure 15 confirms that these patterns remain in regression results.

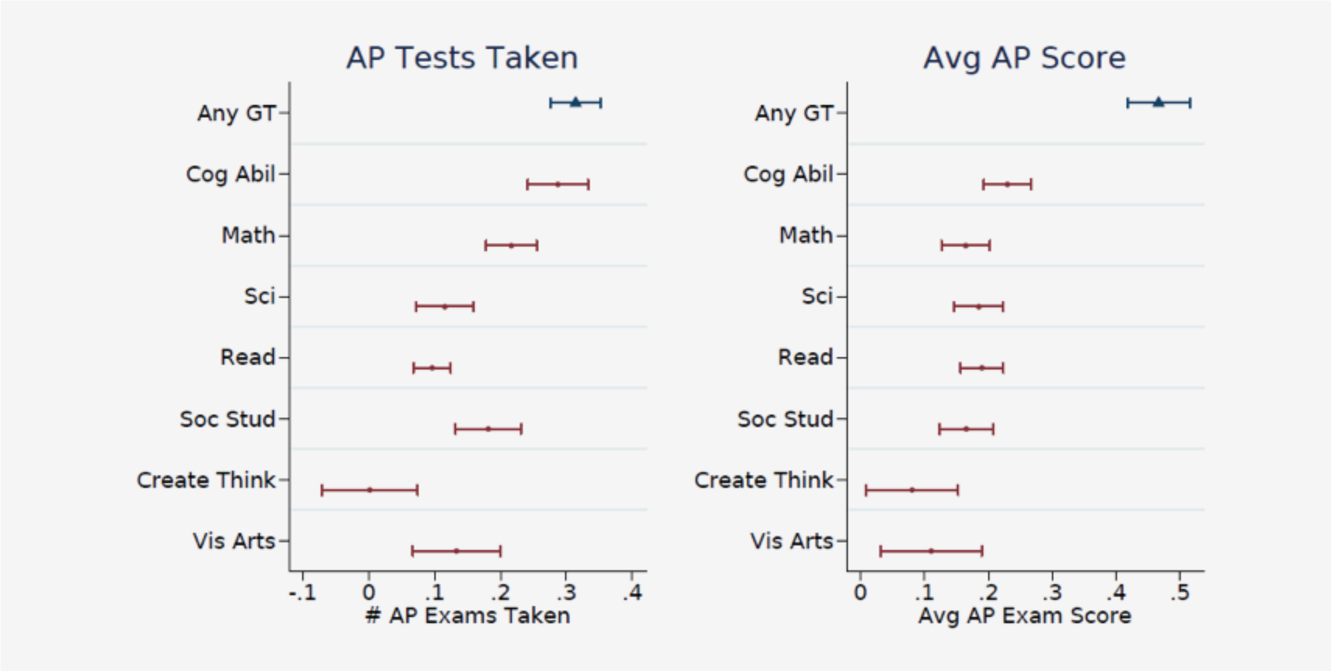
[24] There is some variation, however, by the type of identification. In particular, it is notable that visual or performing arts identification has a relatively small (and often statistically insignificant) relationship with ACT and SAT scores. Further, as expected, students who are identified as gifted in cognitive ability or math perform better than other GT students on the math exams, while those identified in reading perform better (albeit, marginally and not always statistically significantly) than other GT students on the reading exams.

Figure 14: AP exams for high achievers in third-grade-math by GT status (top panel, average number of AP exams taken; bottom panel, average score)



Note: Sample includes all students in 2006–08 third-grade cohorts who scored in the top quintile of the third-grade exam in math and are observed in an Ohio high school. Figure shows means.

Figure 15: Estimates of GT differences in AP testing with demographic controls for third-grade high achievers



Note: Restricted to students in 2006–08 third-grade cohorts who score in the top quintile of math or reading in third grade. Category of GT identification is listed on the y-axis. Students can be identified in multiple categories. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects. The estimate for "any GT" is from a separate regression.

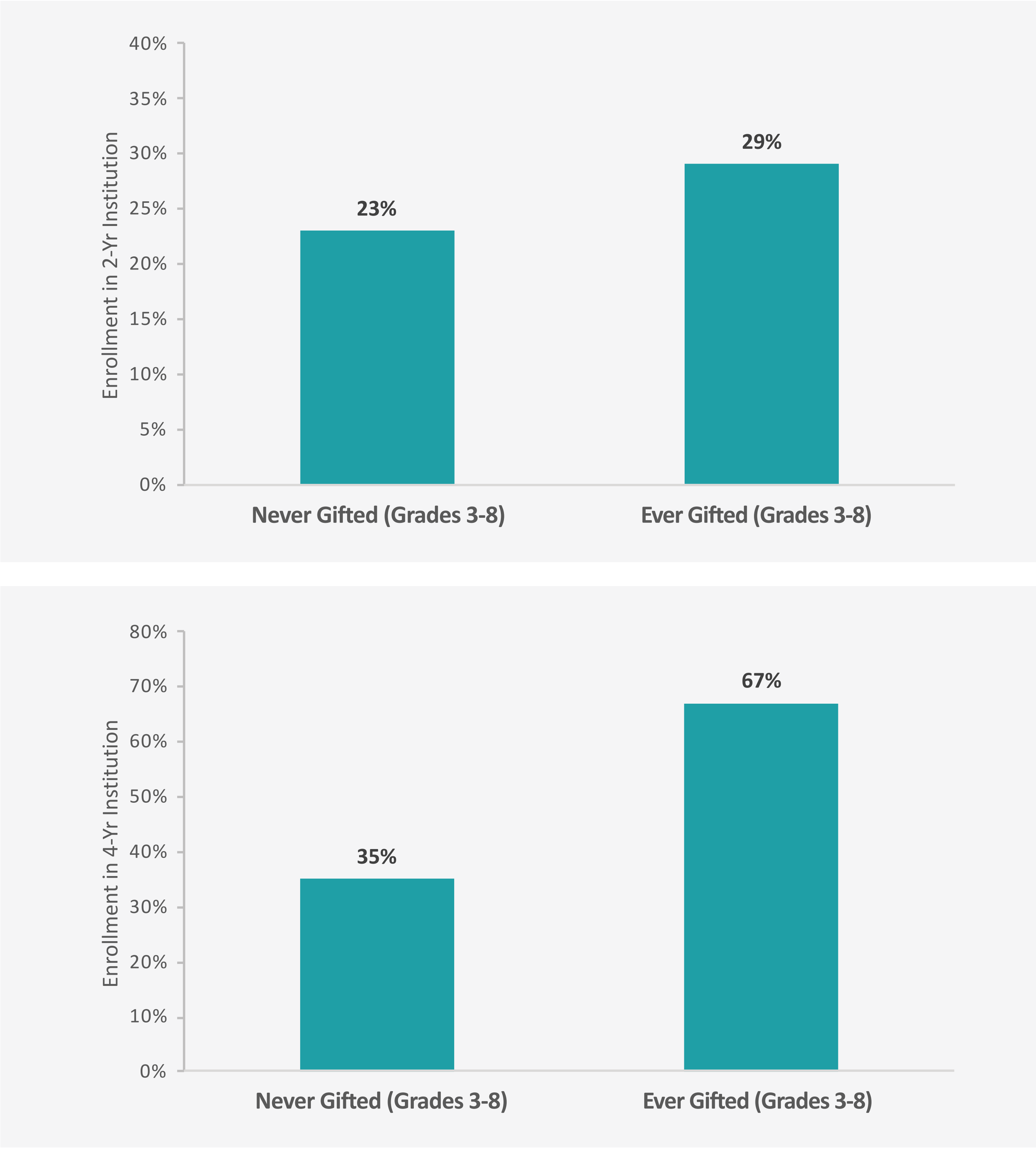
College Enrollment

Finding 12: High achievers who are identified as gifted are substantially more likely to attend four-year colleges, but identified students with similar demographics who attend the same elementary school are only slightly more likely to attend a two-year college.

Starting in Figure 16, we turn to college-going outcomes using NSC data for early high achievers in math (high achievers in reading are discussed in Figure A15) by whether they had any GT identification in grades 3 through 8. Note, once again, that we are restricted to a single cohort for college outcomes. The figure shows that early high achievers with GT identification have higher enrollment in both two-year and four-year schools than non-GT early high achievers. That said, in Figure 17, I provide the regression results. The overall relationship remains, but when looking at specific GT identification categories, the estimates are smaller. For two-year schools in particular, the association between enrollment and any GT is small, while for many specific categories, there is no statistically significant relationship.[25]

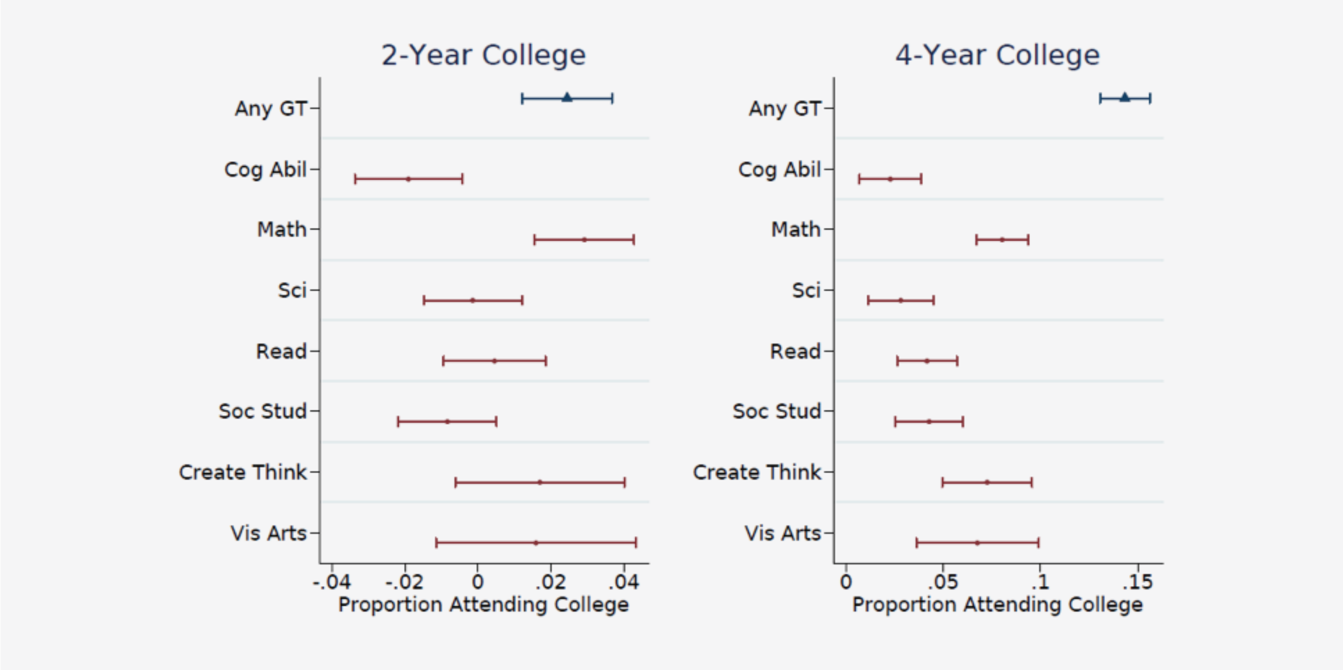
[25] In Figures A16–A18, I provide college-quality measures for math high achievers who attend four-year institutions. A description of the college-quality results is provided in Appendix B.

Figure 16: College enrollment for high achievers in third-grade math by GT status (top panel, enrollment in two-year schools; bottom panel, enrollment in four-year schools)



Note: Restricted to the 2006 grade 3 cohort. Figure shows means.

Figure 17: Estimates of GT gaps in college enrollment with demographic controls for third-grade high achievers



Note: Restricted to students in 2006 third-grade cohort who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

In sum, third-grade high achievers who are identified as GT generally perform better on a range of outcomes than non-GT students, even when accounting for early achievement, demographics, and having attended the same elementary school. The problem is that factors related to GT identification beyond test scores—such as parental persistence, teacher quality, and economic factors—are likely also related to long-term outcomes. Hence, these estimates are not necessarily causal. Nonetheless, they at least do not appear to rule out the potential that GT provides long-term benefits. Unfortunately, there is no way in the existing Ohio data to causally identify these long-term outcomes. However, the panel structure of the data allows me to consider a method that, under certain key assumptions, would enable me to establish the causal impact of GT identification on short-term achievement and absence outcomes, which I turn to in the next section.

Outcomes for Black early high achievers by gifted status

Differences between GT identified and non-GT among Black students are generally similar to what we see in the overall patterns. These are shown in Figures S1 through S4. Although ACT scores, AP exams, and college-enrollment rates are lower for high-achieving Black students than the overall sample, the gap between GT- and non-GT-identified Black students are similar along these metrics. The key exception is high school achievement (OGT) exam scores, where there are notably larger gaps between GT and non-GT Black early high achievers than the overall sample. In particular, math achievement in high school is 1.2 SDs higher for GT-identified Black high achievers than those who are nonidentified, while for reading the gap is 0.8 SDs.

Figure S1: OGT scores for Black high achievers by GT status

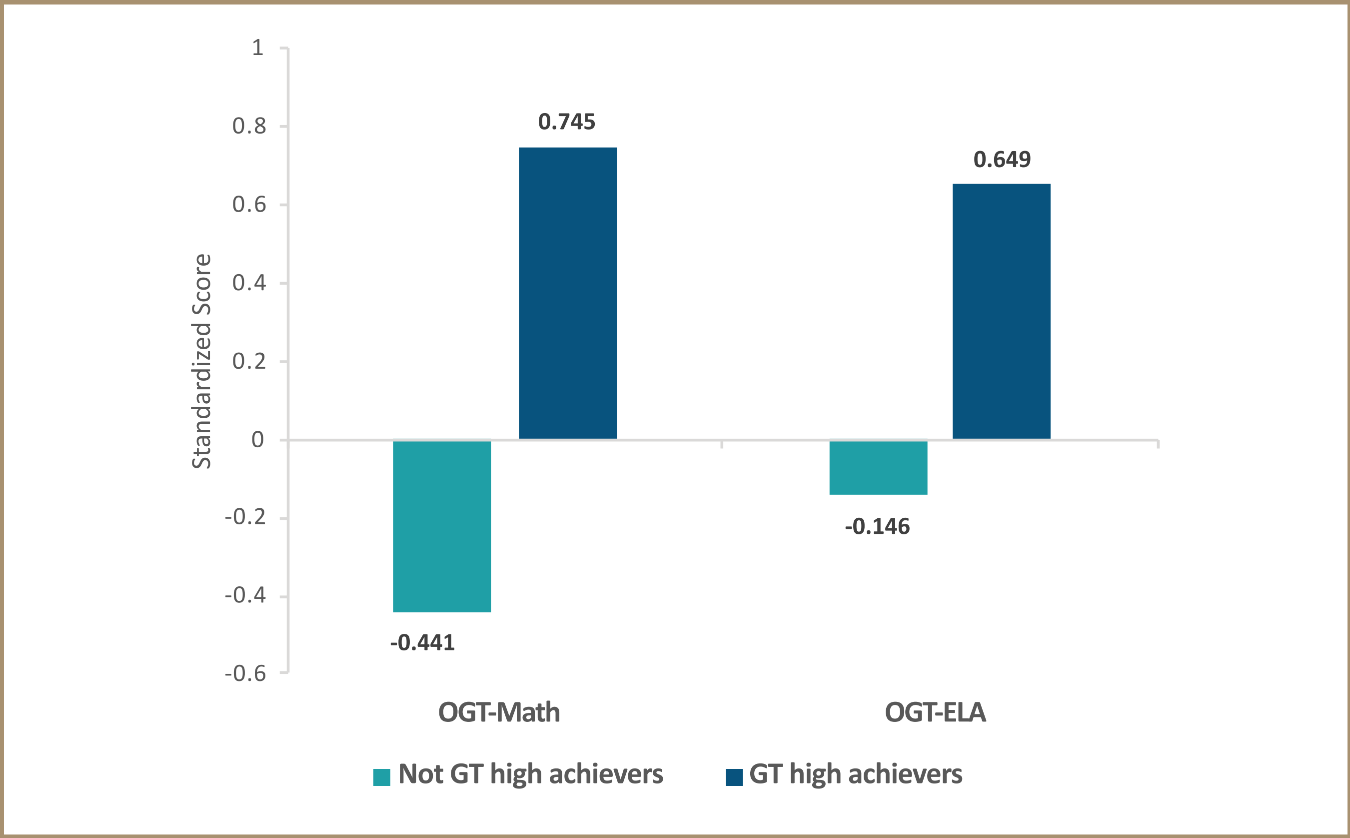
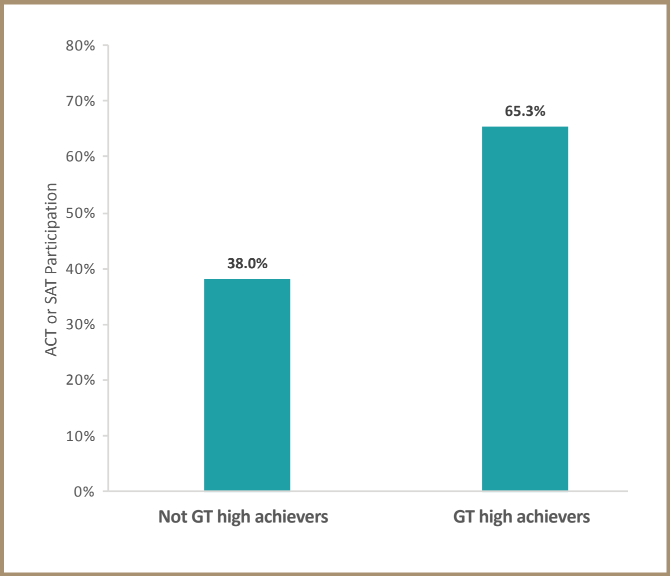


Figure S2: ACT or SAT participation (left) and average ACT scores (right) for Black high achievers by GT status



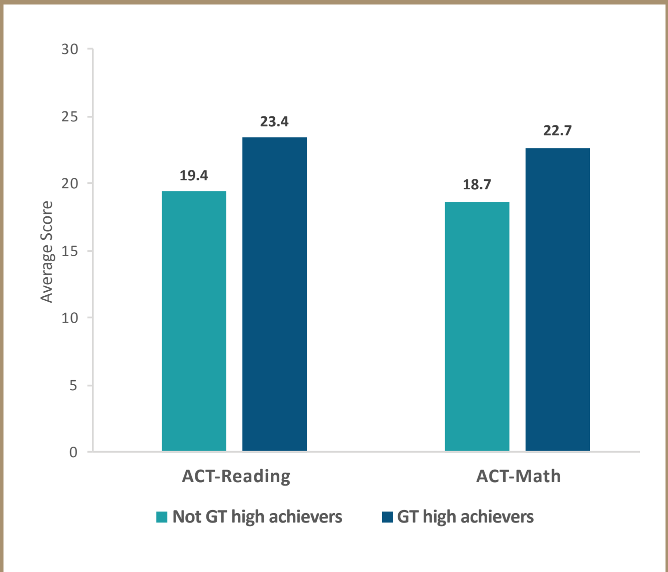


Figure S3: AP-exam participation (left) and average AP score (right) for Black high achievers by GT status

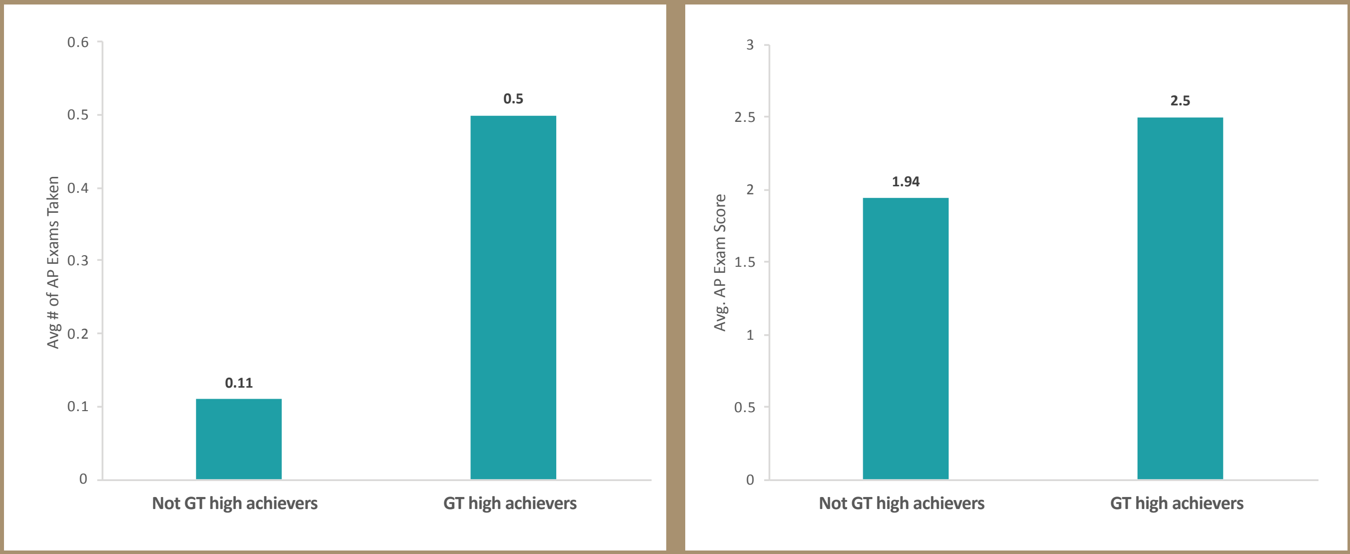
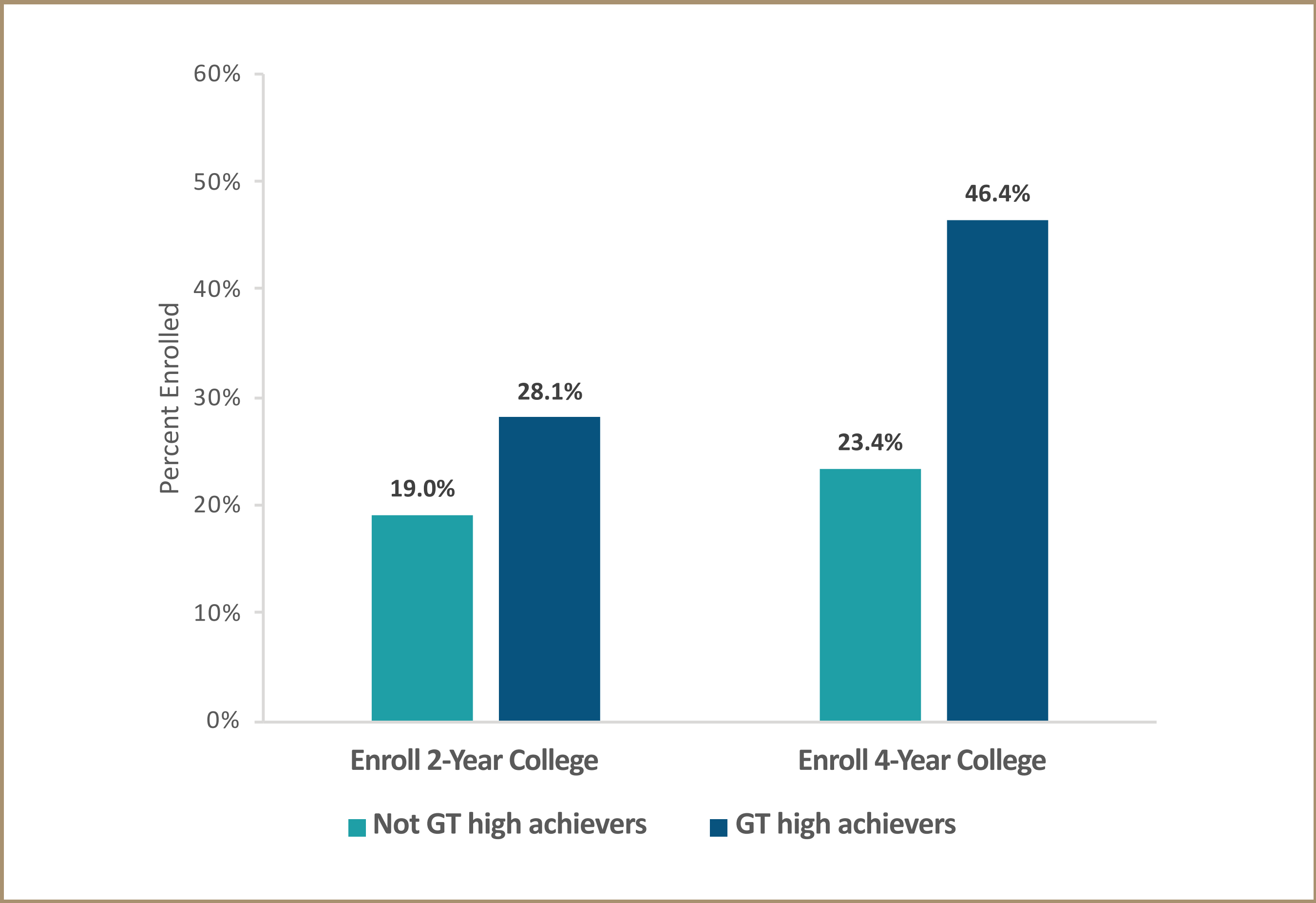


Figure S4: College enrollment for Black high achievers by GT status



Estimates of the impact of gifted identification on achievement

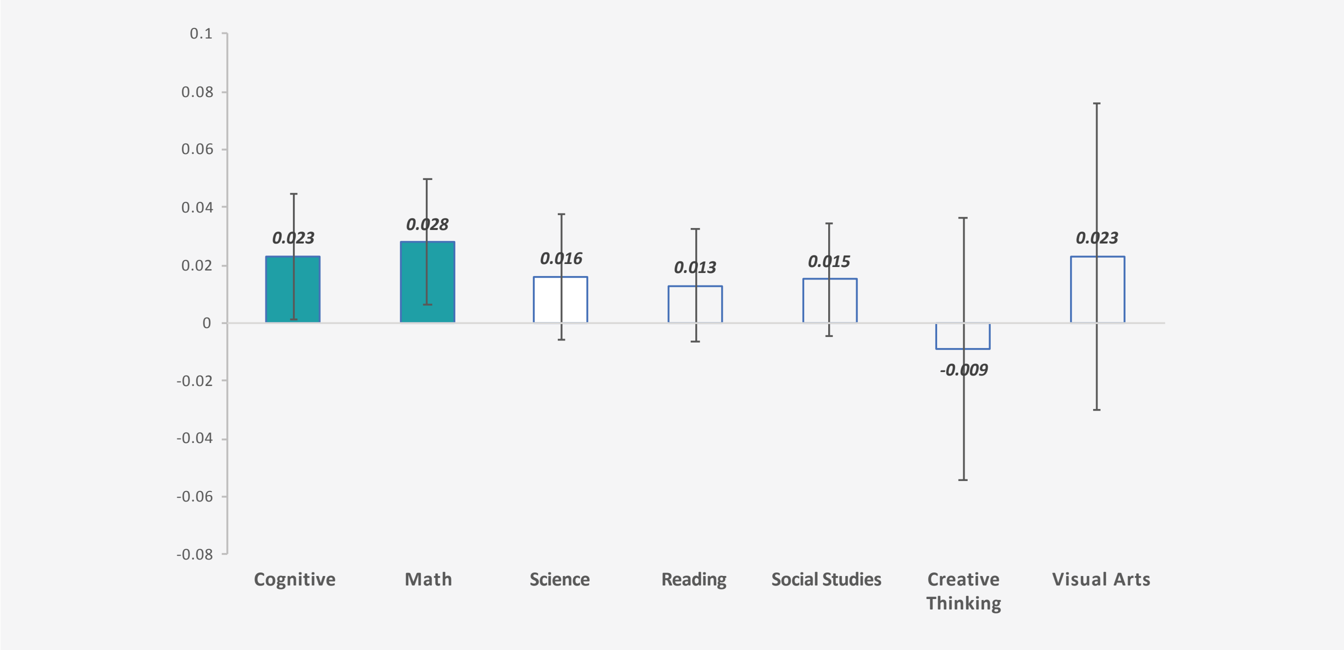
Research question: Does gifted identification improve later achievement on state exams for third-grade high achievers?

Finding 13: Being identified as gifted in math leads to a 0.03 SD increase in annual math scores and a 0.02 SD increase in reading scores. Being identified as gifted in cognitive skills leads to a 0.02 SD increase in annual math scores. No other type of identification generates a statistically significant effect on later achievement.

Finally, we focus on student achievement in grades 4 through 8 and estimate the impact of GT identification. To do this, I rely on a methodology called “student fixed effects.” Details on the method are provided in the Methods section above, and technical details are provided in Appendix A. Ultimately, if we assume that there are no factors that vary over time that affect both gifted identification and achievement, then this model will provide a causal impact of the former on the latter. It should be cautioned that this is a relatively strong assumption, so it remains possible that the true impact of gifted identification on achievement differs.

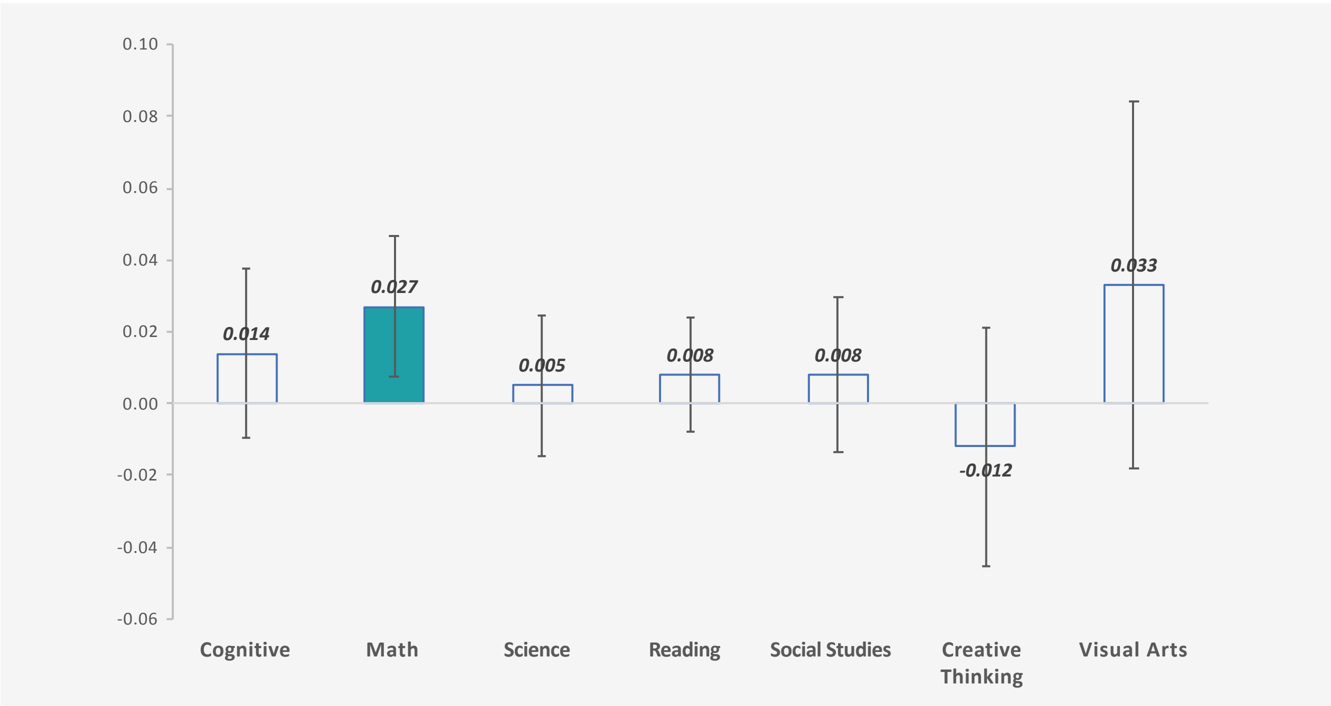
Figure 18 provides the first set of estimates for where I restrict to third-grade high achievers in math and look at how gifted identification affects achievement on state exams. The figure shows regression estimates for each category for identification. The “whiskers” show the 95 percent confidence interval, as in the previous regression figures. First, we look at math and reading achievement. Though most GT classifications have no statistically significant effect on achievement, being identified as gifted in cognitive skills or math significantly increases achievement, by 0.02 to 0.03 SDs per year. Note that when only math identification is included (all other identifications are left out of the model), the effect is larger, at 0.04 SD. Although this may seem small at first glance, we must remember that this only measures the impact of *identification* for gifted services on achievement, rather than the actual receipt of services. During the time frame covered in the data, Ohio was at its nadir in share of identified students receiving services, ranging from 20 percent to 30 percent. Hence, if we assume that all the gifted-identification effects work through service provision, then the impact of receiving services would be approximately four times the estimate (e.g., only one-fourth actually receive services), providing an achievement impact for students receiving services of 0.08 to 0.12 SD in a year.

Figure 18: Impact of gifted identification on math scores, by area of GT classification



The next figure, Figure 19, shows similar results for the reading/ELA exam. Somewhat surprisingly, students who are identified in math show statistically significant achievement impacts of 0.03 SD on reading scores, but the impact of reading identification is insignificant and close to zero. None of the other categories of identification show a statistically significant impact. It is unclear why math identification appears to improve reading scores while reading (and other) identifications do not. It is nonetheless worth noting, though not shown here, that if we only include reading identification without the other categories in the regression, there is a statistically significant impact of 0.02 SD.[26]

Figure 19: Impact of gifted identification on ELA scores, by area of GT classification



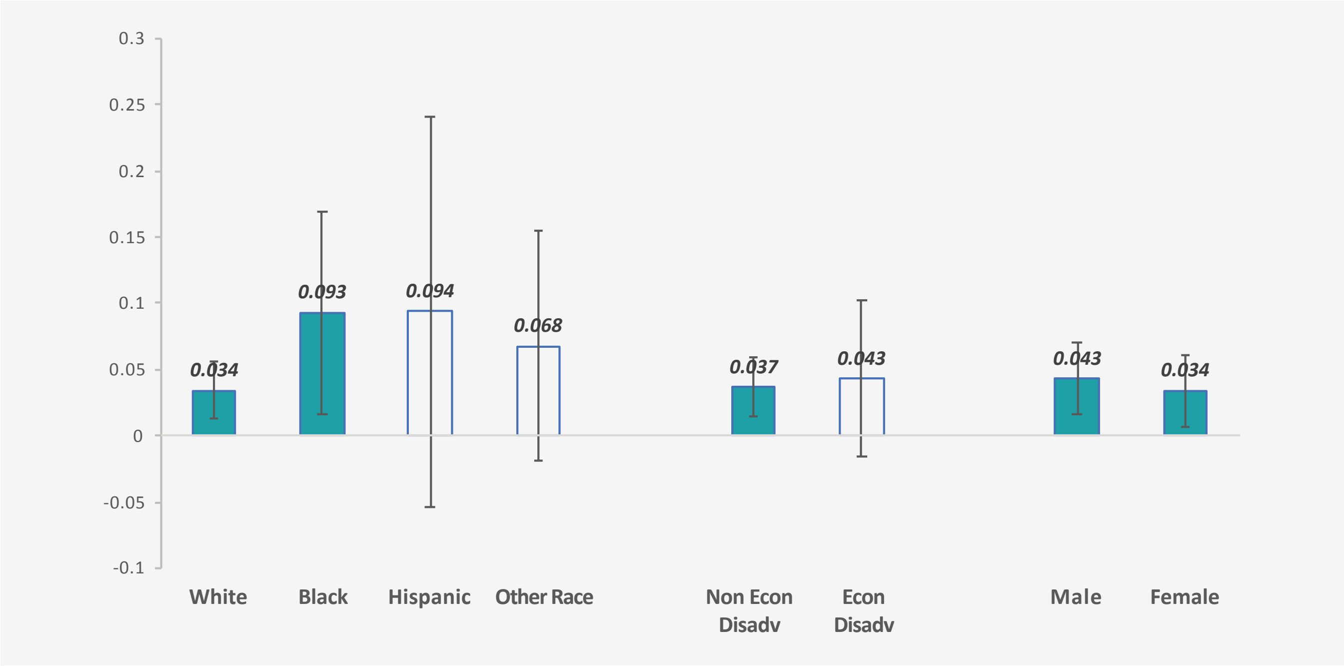
Finding 14: Amongst early high achievers, gifted identification in math generates larger increases in math achievement for Black and Hispanic students than White students. Identification in reading improves reading scores for boys but not girls. Further, schools with higher minority populations show larger increases in both math and reading achievement from GT identification.

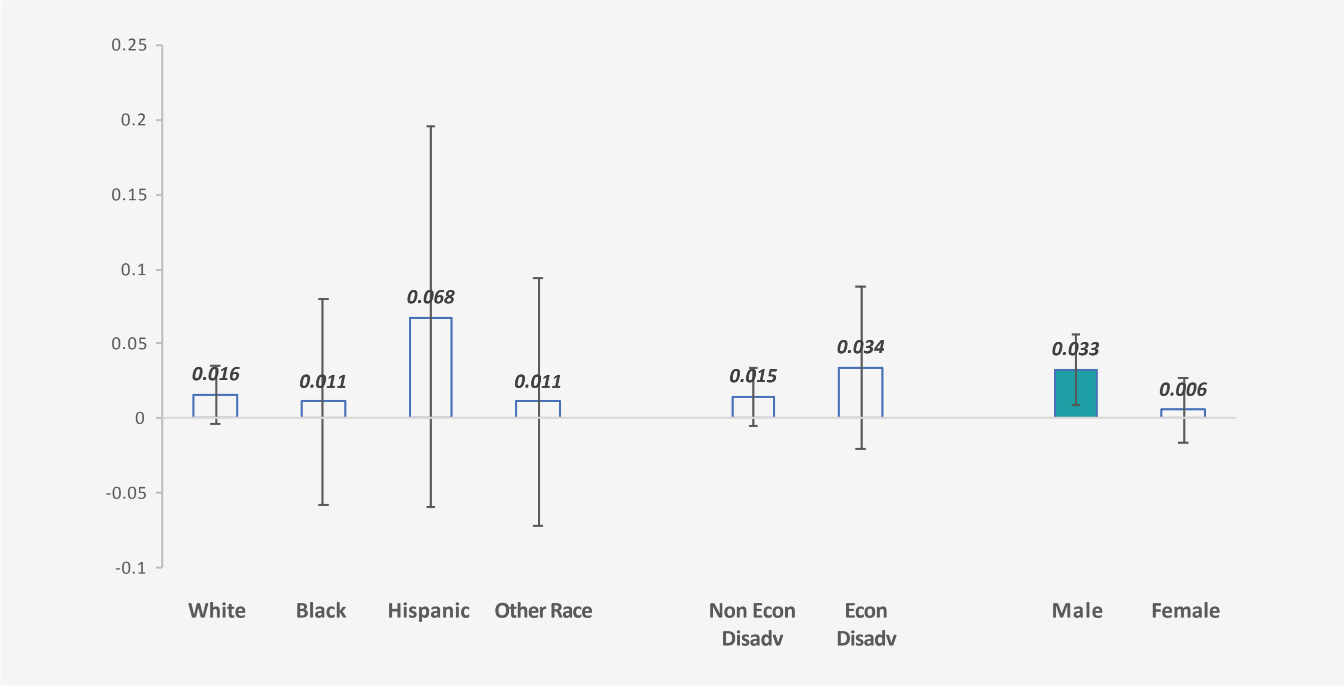
There may be effects for certain groups that do not appear in the overall estimates. Hence, in Figure 20, I provide estimates from student fixed-effects regressions of gifted identification in math on math scores and reading on reading scores.[27] There are a few differences between these estimates that are important. First, minorities appear to benefit more from math identification than White students. Though the impact on high-achieving White students is 0.03 SD, for Black students, it is three times as large (similarly for Hispanic, but the smaller sample size leads to larger standard errors). This is consistent with findings by Card and Giuliano (2016a) that gifted education in a large urban district led to higher performance, particularly amongst minority students. Second, although the reading impacts are relatively similar across races, there is a substantial gender gap in the effects of reading identification—high-achieving boys see an increase in reading achievement of 0.03 SD when they receive a gifted identification in reading, but there is essentially no impact on girls.

26 In Table A2, I provide estimates of the effect of identification in various categories on absences. One argument for the benefits of GT education is that it makes school more interesting for these students and increases their motivation to learn. Although the data are not available to measure this directly, we can proxy for motivation using absence data. Of particular interest is unexcused absences, where presumably the student is out of school for reasons other than illness or family emergencies. Indeed, the estimates are consistently negative (fewer absences) across all identification categories. Nonetheless, the estimates are small. An identified student attends school for 0.06 to 0.08 more days per year depending on the identification category. Even if we assume the effect is driven entirely by service provision, this only increases to at most 0.24 days per year—nonetheless, the figure is statistically significant for some of the estimates.

27 Table A2 also provides estimates with excused and unexcused absences as outcomes. In general, any impacts are small and statistically insignificant.

Figure 20: Impact of GT identification by student demographic (top panel, impact on math scores; bottom panel, impacts on ELA scores)

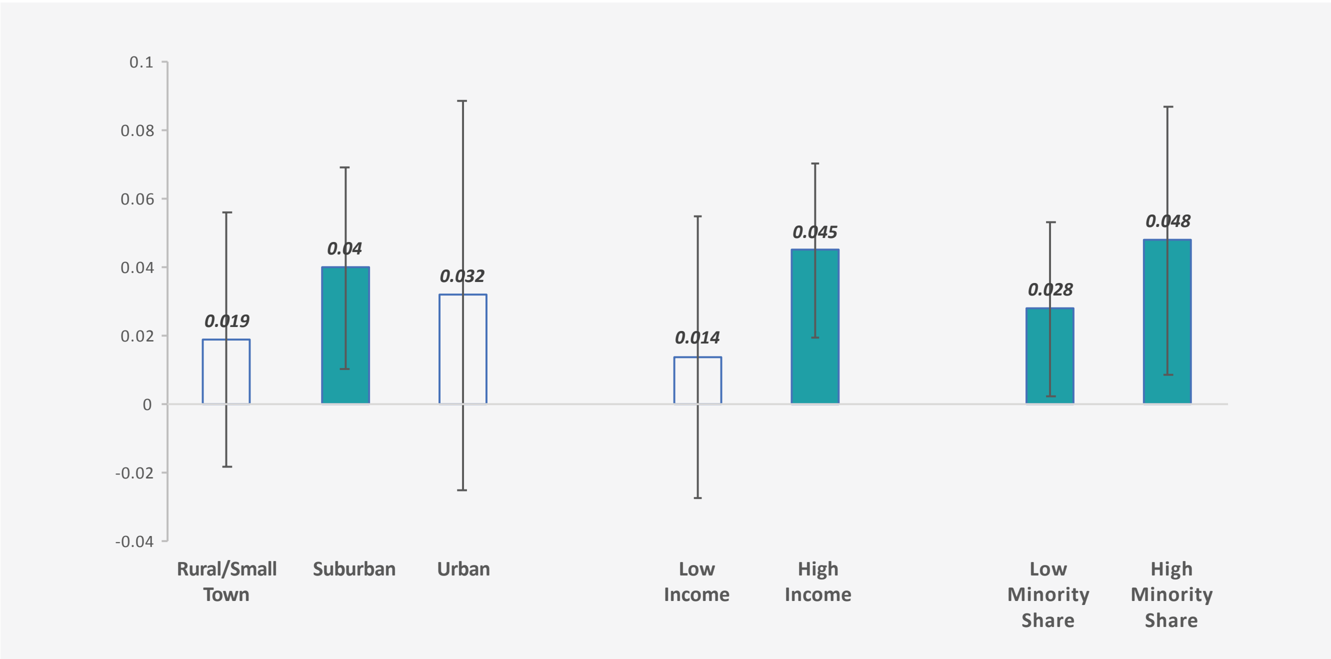


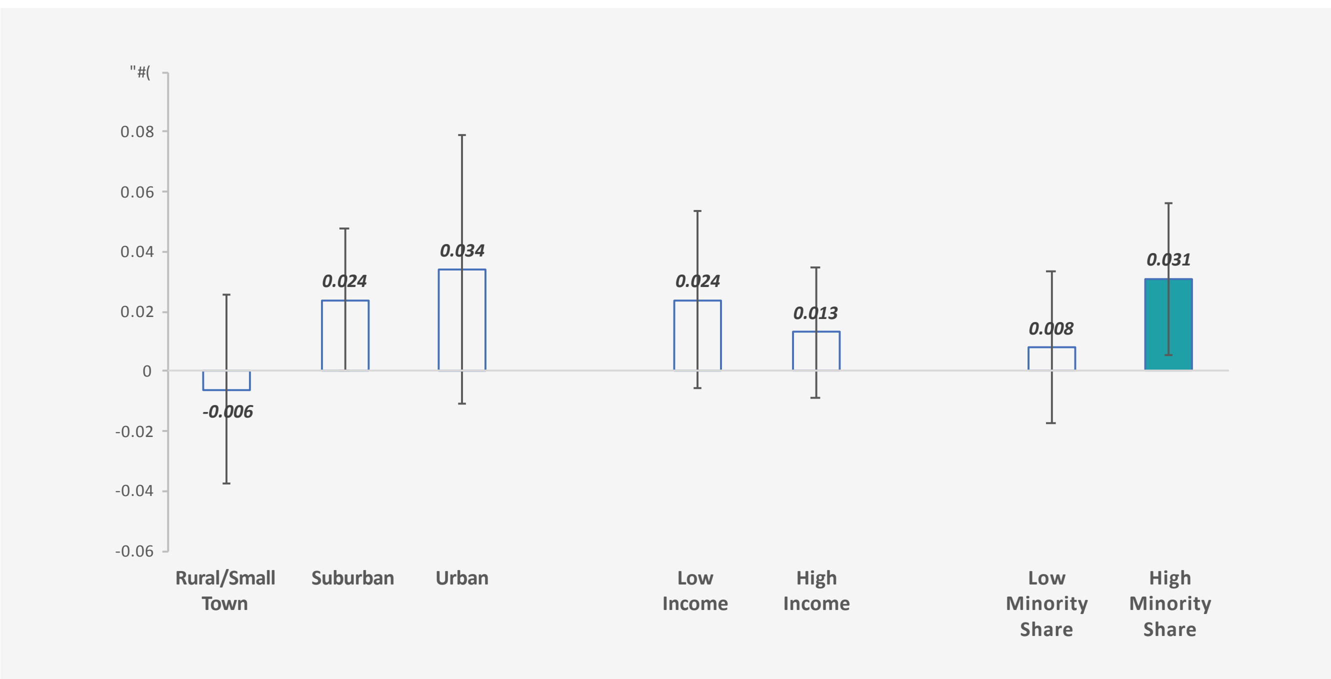


Finally, I turn to Figure 21, which looks at how GT identification affects achievement of early high achievers by school characteristics.[28] Gifted identification in math has larger impacts on students in suburban relative to rural, below-median economic-disadvantage rate (high income), and above-median non-White share (high minority share) schools. For reading, we see similar direction, albeit smaller differences, across urbanicity and minority share of schools. Nonetheless, none of these differences are statistically significant. Thus, in general, a key takeaway is that the results suggest students in more heavily minority but lower-poverty schools benefit more from GT identification.

[28] Table A3 provides estimates for absences by school characteristic. The estimates are again small and mostly statistically insignificant.

Figure 21: Impact of GT identification by school characteristic (top panel, impacts on math scores; bottom panel, impacts on ELA scores)





Conclusion

This study looks at what happens to Ohio's early high achievers (top quintile in math or reading achievement in third grade) as they progress through school. In general, students who perform well in third grade tend to see their achievement fall slightly over time compared to low achievers. More importantly, the study shows that even when students have very similar starting points in third grade, we see substantial differences in long-run outcomes according to their race, gender, and economic status. In particular, early high achieving students who are disadvantaged, Black, or Hispanic tend to perform worse than nondisadvantaged and White students, respectively, along multiple dimensions—high school achievement, college-entrance exams, and AP exams. I also look at college attendance and quality measures of the colleges attended. In these cases, the racial/economic differences are less consistent once one controls for the student's elementary school, cohort, and third-grade achievement. For gender, the gaps that emerge differ across outcomes in whether males, females, or neither perform better. It is important to note that none of these patterns is causal. This part of the study merely shows descriptive trends that highlight potential problems and does not attempt to say whether these demographic characteristics are, in fact, the causes of these gaps.

After establishing these facts, this report went on to look at the potential role that GT identification of students may play in these outcomes. I find, as expected, that GT-identified students perform better than non-GT students who are also high achievers on all the long-run outcomes described above. Further, Black and Hispanic high achievers are less likely to be identified for gifted services relative to White and API students. Similarly, economically disadvantaged students are less likely to be identified than nondisadvantaged students.

To see whether GT identification may have a causal impact on student performance, I rely on a method that provides causal effects if GT identification is not driven by unobserved factors that change over time. The nature of this method limits to contemporaneous rather than long-term outcomes, so I only look at achievement and absences from school. I find small but meaningful impacts of GT identification on math achievement and absences. For absences, the GT impacts are broad, reducing unexcused absences by 0.05 to 0.08 days per year regardless of identification category. Overall, achievement in math and reading increases by 0.02 to 0.03 SDs (approximately one percentile) annually when a student is identified for GT. However, during the time studied, only 20 to 30 percent of GT-identified students were provided services, hence the impact of service provision could be as high as 0.10 to 0.15 SDs per year.

Black students are impacted by GT identification in math more than White students. While the former show improvements of 0.09 SDs annually, the latter only show improvements of 0.03 SDs annually, though these differences are not statistically significant. Hispanic estimates are similar to those for Black students but imprecisely measured due to small sample sizes. Notably, this is consistent with the results in Card and Giuliano (2016a) who found little impact of GT on White students but large impacts on Black and Hispanic students. Hence, the results in both of our studies suggest that GT education could be helpful in improving outcomes for high-achieving Black or Hispanic students. I also find that students from schools with high minority shares show larger, albeit not statistically different, impacts from GT identification, though it is also the case that students in high-income schools benefit more than students in low-income schools.

This study provides some evidence that closing the "GT identification gap" could help to close the "excellence gap." In other words, if early-high-achieving, low-income students and students of color were identified for GT programs at the same rate as their White and Asian high-achieving peers, this could reduce achievement gaps down the road, though it is by no means a panacea as the estimated effects are relatively small. Though the evidence provided here on long-run outcomes needs far stronger assumptions to say that GT identification increases these outcomes, combined with the achievement results, they nonetheless suggest that if these third-grade "high flyers" gain GT identification, they would be more likely to go on to demonstrate higher achievement on state exams, AP exams, and the ACT or SAT—and, as a result, matriculate at higher rates into four-year colleges, including elite institutions.

One potential policy response that could help achieve these goals is to increase the use of universal screening for GT services, which has been shown to increase participation by students from underrepresented minorities (Card and Giuliano 2016b), though it is important that screening mechanisms are nondiscriminatory (Joseph and Ford 2006). Indeed, Ohio did just this starting in 2017, but it remains to be seen how effective universal screening has been for improving access. A common refrain is that “college starts in Kindergarten.” If we want to diversify our institutions of higher education by race and class and if we want to make headway in reducing racial gaps in income and other long-term outcomes, closing the “GT identification gap” for high-achieving students would go a long way.

Appendixes

Appendix A: Technical Appendix

Regression analysis of high school and college outcome gaps for high achievers

For Figures A1, A3, A5, A7, and A10, I estimate the following ordinary least squares (OLS) regression model:

$$Y_{ics} = \beta_0 + \beta_1 Female_i + \beta_2 Black_i + \beta_3 Hispanic_i + \beta_4 API_i + \beta_5 EconDisadv_i + \omega_1 OAAMath_{ig3} + \omega_1 OAARead_{ig3} + \gamma_{cs} + \varepsilon_{ics}$$

Where Y_{ics} is an outcome in high school or college for student *i* in third-grade cohort *c* and high school *s*. $Female_i$ is an indicator that equals 1 if a student is female, $Black_i$ is an indicator that equals 1 if the student is Black, $Hispanic_i$ is an indicator that equals 1 if a student is Hispanic, API_i is an indicator that equals 1 if the student is API, and $EconDisadv_i$ is an indicator that equals 1 if whether the student is eligible for free or reduced-price meals at any time between grades 3 and 12. $\omega_1 OAAMath_{ig3}$ and $\omega_1 OAARead_{ig3}$ are third-grade math and reading achievement scores, respectively. γ_{cs} is a set of cohort-by-school fixed effects based on the student's third-grade school. Standard errors are clustered by the student's school district when enrolled in third grade.

Regression analysis of GT identification gaps in high school and college outcomes for high achievers

For Figures 11, 13, 15, and 17 and Figures A11, A13, and A18, I estimate the following OLS regression model:

$$Y_{ics} = \beta_0 + \beta_1 Female_i + \beta_2 Black_i + \beta_3 Hispanic_i + \beta_4 API_i + \beta_5 EconDisadv_i + \delta_1 GTID_{Cog} + \delta_2 GTID_{Math} + \delta_3 GTID_{Sci} + \delta_4 GTID_{Read} + \delta_5 GTID_{SocStud} + \delta_6 GTID_{Think} + \delta_7 GTID_{VisArt} + \omega_1 OAAMath_{ig3} + \omega_2 OAARead_{ig3} + \gamma_{cs} + \varepsilon_{ics}$$

Hence, in addition to the variables in (1), I add a series of indicator variables for whether the student is identified for GT at any point between grades 3 and 8. These include indicators for identification as gifted in cognitive skills (*Cog*), mathematics (*Math*), science (*Sci*), reading (*Read*), social studies (*SocStud*), creative thinking (*Think*), and visual or performing arts (*VisArts*). These categories are not mutually exclusive, so students can qualify under multiple categories. Standard errors are clustered by the student's school district when enrolled in third grade.

Individual fixed-effects regression analysis of effect of GT identification on achievement and attendance

To estimate the impacts of GT identification on student achievement (Figures 18–21 and Tables A2 and A3), I utilize an individual (student) fixed-effects strategy. This model uses changes in achievement before and after they are identified (in Ohio, identification cannot be removed) to identify the causal effect of GT on student performance. It relies on the assumption that any unobserved variables that may affect both GT identification and student achievement independently do not, on their own, change over time. Specifically, I estimate the following econometric model:

$$Y_{icgs} = \beta_0 + \sum_{GT} \beta_{GT} Gifted^{GT}_{icg} + \gamma EconDisadv_{icg} + \delta_i + \gamma_{gc} + \omega_s + \tau_t + \varepsilon_{ics}$$

where Y_{icgs} is achievement of student *i* in cohort *c*, grade *g*, and school *s*; $Gifted^{GT}_{icg}$ is an indicator that equals 1 if the student is identified as gifted in a given category *GT ε (Cognitive Ability, Mathematics, Science, Reading,* *Social Studies, Creative Thinking, Visual or Performing Arts)* and 0 otherwise; $EconDisadv_{icg}$ is as defined in part A of the technical appendix; and δ_i, γ_{gc}, ω_s, and τ_t are individual, grade-by-cohort, school, and year fixed effects, respectively.

I also estimate models that use only a single GT identification category as follows:

$$Y^{Math}_{icgs} = \beta_0 + \beta_0 Gifted^{Math}_{icg} + \gamma EconDisadv_{icg} + \delta_i + \gamma_{gc} + \omega_s + \tau_t + \varepsilon_{ics}$$

$$Y^{Read/ELA}_{icgs} = \beta_0 + \beta_0 Giftedi^{Read}_{cg} + \gamma EconDisadv_{icg} + \delta_i + \gamma_{gc} + \omega_s + \tau_t + \varepsilon_{ics}$$

$$Y^{Absences}_{icgs} = \beta_0 + \beta_0 Gifted^{Any}_{icg} + \gamma EconDisadv_{icg} + \delta_i + \gamma_{gc} + \omega_s + \tau_t + \varepsilon_{ics}$$

where $Gifted^{Any}_{icg}$ indicates the student was identified as gifted in any underlying category. Standard errors for all these models are clustered by the student's school district when enrolled in third grade.

Variable descriptions

Variable	Description
ACT Math	Score (out of 36) on ACT Math exam.
ACT Reading	Score (out of 36) on ACT Reading exam.
AP tests taken	Total number of AP exams taken by the student throughout high school.
API	Student's race/ethnicity identified as API.
Average AP score	The mean score for a student of AP exams taken. Exams are on a 1 to 5 scoring scale.
Black	Student's race/ethnicity identified as Black.
Economically disadvantaged	Student qualifies for free or reduced-price lunch.
Female	Student's gender listed as female.
Four-year-college enrollment	Whether student enrolls in a program leading to a bachelor's degree within two years after completing high school.
GT status	Whether a student is identified as GT. For long-term outcomes, this is identification at any time between grades 3 and 8. For pre–high school achievement outcomes, this is identification in the current year. Note that identification for GT services does not mean that a student receives services—only that he or she is eligible if services are offered by the school.
GT – cog ability	Student identified as being gifted in cognitive ability.
GT – creative thinking	Student identified as being gifted in creative thinking ability.
GT – math	Student identified as being gifted in math ability.
GT – read	Student identified as being gifted in reading ability.
GT – science	Student identified as being gifted in science ability.
GT – social studies	Student identified as being gifted in social studies ability.
GT – visual or performing arts	Student identified as being gifted in visual or performing arts ability.
Hispanic	Student's race/ethnicity identified as non-Black Hispanic.
Immigrant	Student is an immigrant.
Limited English proficiency	Student is an English language learner or has limited English proficiency.
OAA achievement	Score on OAA exam. Expressed as SDs from the mean score in each grade and year unless specified otherwise. Scores provided for math and reading exams in grades 3 through 8.
Took SAT/ACT	Student has taken the SAT or ACT at some point while enrolled in high school.
Two-year-college enrollment	Whether student enrolls in a program leading to an associate degree within two years after completing high school.
White	Student's race/ethnicity identified as White/Caucasian.

Appendix B: College-quality results

In addition to the college-enrollment data, each college in the NSC data was matched by CHRR to data from the Integrated Postsecondary Education Survey (IPEDS) to incorporate measures of college quality. Understanding how high achievers of different demographics fare in terms of getting access to higher-quality and elite institutions is important as society becomes more concerned about a lack of diversity in the highest reaches of the postsecondary education system. To measure college quality, the study uses the seventy-fifth percentiles of SAT and ACT scores of incoming students for each institution, as reported to IPEDS.

In addition to college enrollment, I look at differences in the quality of college attended. Due to data security restrictions, the exact information for each school from IPEDS could not be matched directly to the student-level data. So instead, institutions that report their seventy-fifth percentiles of SAT or ACT scores for their freshmen classes are placed into deciles of institutions. Those that do not report these scores I consider to be "noncompetitive." For example, a value for SATs of six indicates that the student attended a college that was in the sixtieth to seventieth percentile of all colleges in terms of the seventy-fifth percentile of entering-student SAT scores at that college. Note that the noncompetitive category includes both two-year and four-year institutions, while the ranked schools include only four-year institutions. Interpretation of these measures must be done cautiously, given that one must restrict the data to students who attend an NSC-matched institution in the first place. With that in mind, given leaving people who do not attend these institutions out of the analysis likely leads to understating the differences in college quality between high- and low-achieving students, it is notable that in Panel C of Table A1, we still see clear increases along all three quality measures as third-grade achievement increases.

Within the four-year institutions we can consider college selectivity. Multiple recent studies have suggested that the selectivity of the institution a student attends is important for their outcomes (Long 2008; Black and Smith 2006), particularly for minorities (Dale and Krueger 2014), and many high-achieving students do not attend higher-quality institutions even when they are a good fit and the costs to the students are low (Dynarski et al. 2018; Hoxby and Turner 2015; Hoxby and Avery 2013). These measures are restricted to students who attend four-year institutions. Figures A8 and A9 show box plots for how each college that each student attends ranks relative to others in the seventy-fifth SAT or ACT percentile of their students. In general, conditional on attending a four-year school, there is little variation in mean enrollments—based on ACTs, the average selectivity decile is six across all subgroups—though the spread of institutional selectivity is wider for nondisadvantaged relative to disadvantaged and minority students relative to White students. Hence, overall, there is little difference in the quality of college attended by high-achieving students from these groups, given they attend a four-year college.

Turning to regression results in Figure A10, the lack of precision makes it difficult to tell whether the estimates are statistically different. Nonetheless, in the bottom panel, we see statistically significantly lower enrollment in nonselective institutions by women relative to men and by API students relative to White students. Hispanic students, however, are more likely to attend lower-selectivity institutions, conditional on attending any four-year college.

Appendix C: Supplementary tables and figures

Table A1: Means and SDs of important variables

Panel A: Demographics in grade 3 by third-grade math quintile

Variable	All students (1)	Lowest quintile (2)	2nd quintile (3)	3rd quintile (4)	4th quintile (5)	Highest quintile (6)
White	0.742	0.539	0.706	0.781	0.829	0.861
	(0.437)	(0.498)	(0.456)	(0.414)	(0.377)	(0.346)
Black	0.158	0.342	0.188	0.123	0.083	0.049
	(0.364)	(0.474)	(0.391)	(0.329)	(0.276)	(0.215)
Hispanic	0.039	0.059	0.046	0.038	0.030	0.021
	(0.193)	(0.236)	(0.210)	(0.190)	(0.169)	(0.144)
Other	0.056	0.067	0.057	0.051	0.048	0.056
	(0.230)	(0.250)	(0.232)	(0.221)	(0.214)	(0.230)
Female	0.489	0.488	0.501	0.495	0.486	0.473
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.499)
Immigrant	0.009	0.015	0.009	0.007	0.007	0.008
	(0.096)	(0.123)	(0.093)	(0.086)	(0.083)	(0.087)
Limited English proficiency	0.027	0.047	0.030	0.024	0.018	0.015
	(0.161)	(0.211)	(0.170)	(0.152)	(0.133)	(0.122)
Economically disadvantaged	0.230	0.361	0.275	0.222	0.173	0.117
	(0.421)	(0.480)	(0.446)	(0.416)	(0.378)	(0.321)
Maximum observations	901175	181423	183904	174975	183565	177308

Panel B: Standardized achievement

Variable	All students	Lowest quintile	2nd quintile	3rd quintile	4th quintile	Highest quintile
Math by third-grade math quintile						
Grade 3 math	0.003	-1.386	-0.453	0.037	0.504	1.343
	(0.999)	(0.675)	(0.188)	(0.146)	(0.163)	(0.524)
Observations	900647	181091	183823	174948	183515	177270
Grade 8 math	0.010	-0.872	-0.281	0.113	0.463	0.899
	(0.942)	(0.763)	(0.707)	(0.715)	(0.729)	(0.777)
Observations	571416	119555	126735	123577	114354	87195
ELA/reading by third-grade reading quintile						
Grade 3 reading	0.004	-1.085	-0.262	0.133	0.461	0.890

	(0.996)	(0.925)	(0.663)	(0.647)	(0.669)	(0.731)
Observations	866008	179887	181234	170153	174149	160585
Grade 8 ELA	0.026	-0.872	-0.307	0.046	0.385	0.805
	(0.956)	(0.831)	(0.749)	(0.749)	(0.759)	(0.778)
Observations	334823	63304	69887	66123	68538	66971

Panel C: High School and College Outcomes by Math Quintile

Variable	All students	Lowest quintile	2nd quintile	3rd quintile	4th quintile	Highest quintile
Took ACT or SAT	0.48	0.22	0.38	0.50	0.59	0.73
	(0.50)	(0.41)	(0.48)	(0.50)	(0.49)	(0.44)
Observations	351188	70046	73546	68625	70536	68435
ACT Reading	21.9	16.4	18.9	20.8	22.6	25.2
	(5.5)	(3.9)	(4.3)	(4.6)	(4.9)	(5.1)
Observations	162220	14271	26724	33348	40095	47782
ACT Math	21.4	16.2	18.2	20.2	22.2	25.0
	(4.7)	(2.3)	(3.1)	(3.7)	(4.0)	(4.2)
Observations	162867	14123	26792	33528	40381	48043
ACT Writing	10.3	9.5	9.6	9.8	10.7	10.7
	(6.3)	(5.2)	(5.6)	(5.9)	(6.6)	(6.8)
Observations	85656	6161	12871	17059	20862	28703
SAT Reading	555.05	410.05	471.71	516.79	559.53	612.95
	(107.07)	(72.89)	(85.18)	(85.33)	(88.34)	(89.35)
Observations	20322	1408	2314	3207	4790	8603
SAT Math	563.47	399.20	465.30	519.07	567.30	630.58
	(111.21)	(69.44)	(78.80)	(83.48)	(85.95)	(88.68)
Observations	20313	1381	2308	3211	4806	8607
Took an AP exam	0.11	0.01	0.04	0.09	0.15	0.26
	(0.31)	(0.11)	(0.21)	(0.29)	(0.36)	(0.44)
Observations	351188	70046	73546	68625	70536	68435
Avg AP score	2.74	1.71	2.06	2.37	2.67	3.09
	(1.10)	(0.86)	(1.00)	(0.99)	(1.03)	(1.05)
Observations	38738	912	3236	6368	10596	17626
AP exam count	0.26	0.02	0.08	0.18	0.33	0.70
	(0.96)	(0.21)	(0.43)	(0.72)	(1.01)	(1.62)
Observations	351188	70046	73546	68625	70536	68435
Two-year-college enrollment	0.24	0.16	0.23	0.26	0.28	0.29
	(0.43)	(0.37)	(0.42)	(0.44)	(0.45)	(0.45)
Observations	126814	25175	28353	25226	22165	25895
Four-year-college enrollment	0.37	0.13	0.27	0.39	0.48	0.60
	(0.48)	(0.33)	(0.44)	(0.49)	(0.50)	(0.49)

Observations	126814	25175	28353	25226	22165	25895
Public four-year-college enrollment	0.29	0.09	0.21	0.30	0.38	0.48
	(0.45)	(0.29)	(0.41)	(0.46)	(0.49)	(0.50)
Observations	126814	25175	28353	25226	22165	25895
Private four-year-college enrollment	0.10	0.04	0.08	0.11	0.13	0.17
	(0.31)	(0.19)	(0.27)	(0.31)	(0.34)	(0.38)
Observations	126814	25175	28353	25226	22165	25895
Decile of colleges' 75th	6.01	5.79	5.91	5.92	5.97	6.09
Percentile ACT scores	(2.21)	(2.36)	(2.18)	(2.18)	(2.18)	(2.22)
Observations	6870	203	729	1205	1677	3056
Decile of colleges' 75th	6.38	6.27	6.22	6.29	6.33	6.49
Percentile SAT scores	(2.20)	(2.29)	(2.24)	(2.19)	(2.20)	(2.19)
Observations	6899	203	729	1214	1681	3072
College did not report	0.89	0.97	0.94	0.91	0.87	0.83
ACT/SAT to IPEDS	(0.31)	(0.18)	(0.24)	(0.29)	(0.33)	(0.38)
Observations	63030	6389	11841	13388	13462	17950

Panel D: GT identification by grade 3 achievement

Variable	All students	Lowest quintile	2nd quintile	3rd quintile	4th quintile	Highest quintile
Grade 3						
Math by G3 math test	0.068	0.002	0.008	0.026	0.074	0.233
	(0.252)	(0.047)	(0.091)	(0.158)	(0.262)	(0.423)
Observations	901353	181480	183949	175016	183586	177322
Reading/writing by G3 reading test	0.061	0.003	0.008	0.026	0.073	0.199
	(0.240)	(0.056)	(0.090)	(0.162)	(0.259)	(0.399)
Observations	866773	177178	176781	168882	171617	172315
Any by G3 math test	0.148	0.012	0.039	0.090	0.190	0.414
	(0.355)	(0.110)	(0.193)	(0.287)	(0.392)	(0.493)
Observations	901353	181480	183949	175016	183586	177322
Grade 8						
Math by G3 math test	0.110	0.002	0.011	0.041	0.129	0.364
	(0.312)	(0.047)	(0.103)	(0.199)	(0.335)	(0.481)
	810489	157419	165198	158954	167539	161379
Reading/writing by G3 reading test	0.086	0.003	0.009	0.036	0.103	0.278
	(0.281)	(0.056	(0.098)	(0.186)	(0.305)	(0.448)
	779442	154403	159024	153452	156360	156203
Any by G3 math test	0.216	0.016	0.055	0.137	0.292	0.572
	(0.411)	(0.126)	(0.228)	(0.344)	(0.454)	(0.495)
Observations	810499	157419	165198	159954	167539	161379

Table A2: Effects of GT identification on absences for third-grade high achievers by student characteristics

Characteristic →	White (1)	Black (2)	Hispanic (3)	Other race (4)	Not disadvantaged (5)	Disadvantaged (6)	Male (7)	Female (8)
A. Excused absences								
GT identification in	0.043	0.051	0.470	0.021	0.037	0.261*	-0.035	0.166*
any category	(0.077)	(0.295)	(0.333)	(0.213)	(0.080)	(0.147)	(0.093)	(0.097)
Observations	448923	29453	10242	22043	429168	88511	255791	261900
B. Unexcused absences								
GT identification in	-0.081	-0.117	-0.399*	-0.142	-0.090	-0.224	-0.097	-0.132
any category	(0.057)	(0.276)	(0.238)	(0.119)	(0.059)	(0.139)	(0.061)	(0.082)
Observations	448923	29453	10242	22043	429168	88511	255791	261900

Each regression contains individual and school fixed effects, along with controls for economic disadvantage.

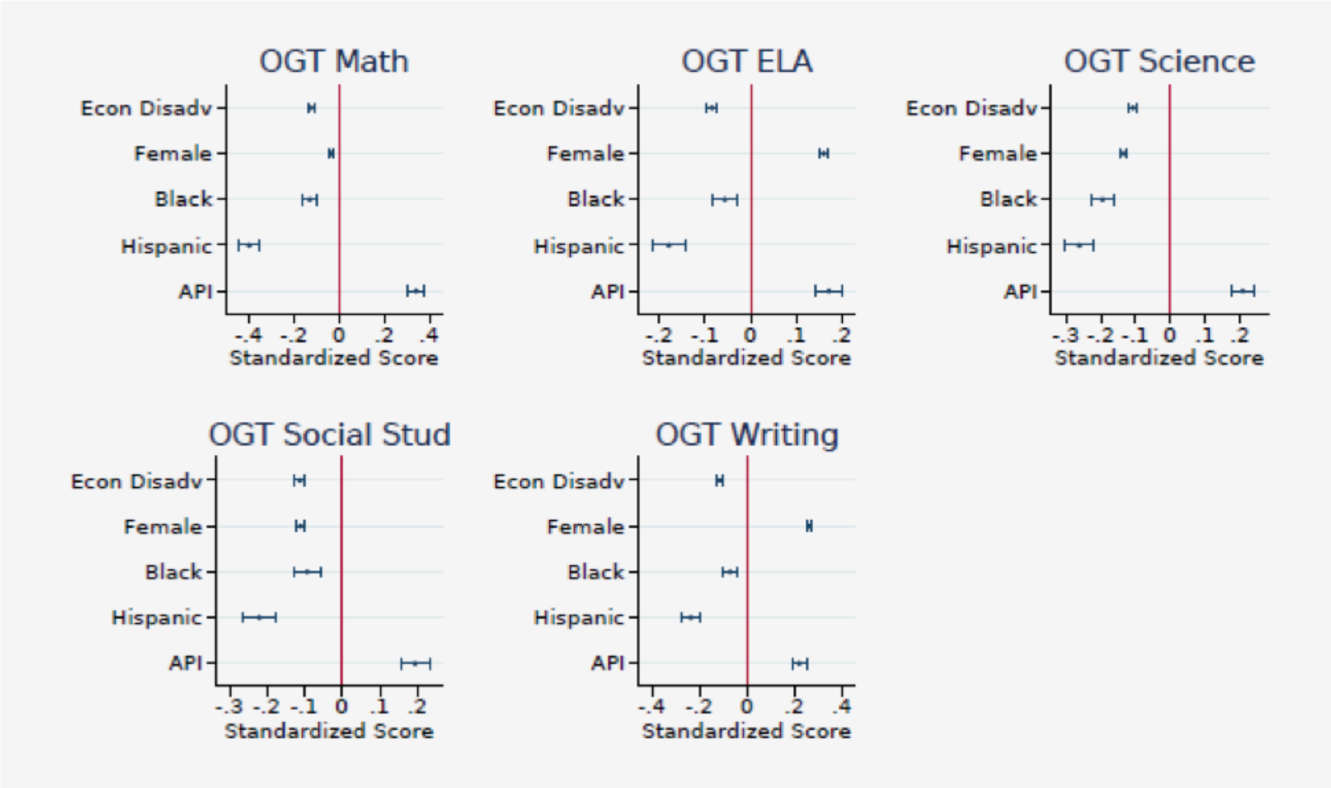
Table A3: Effects of GT identification on absences of third-grade high achievers by school characteristics

Characteristic →	Rural/small town (1)	Suburban (2)	Urban (3)	Low income (4)	High income (5)	Low minority share (6)	High minority share (7)
A. Excused absences							
GT identification in	0.096	-0.106	0.352	0.241	-0.066	-0.004	0.128
any category	(0.082)	(0.084)	(0.330)	(0.163)	(0.073)	(0.067)	(0.175)
Observations	202158	221379	70438	184187	310227	306392	188622
B. Unexcused absences							
GT identification in	-0.075*	-0.027	-0.416	-0.210	-0.061*	-0.059**	-0.20
any category	(0.041)	(0.033)	(0.293)	(0.158)	(0.034)	(0.025)	(0.152)
Observations	202158	221379	70438	184187	310227	306392	188622

Each regression contains individual and school fixed effects, along with controls for economic disadvantage.

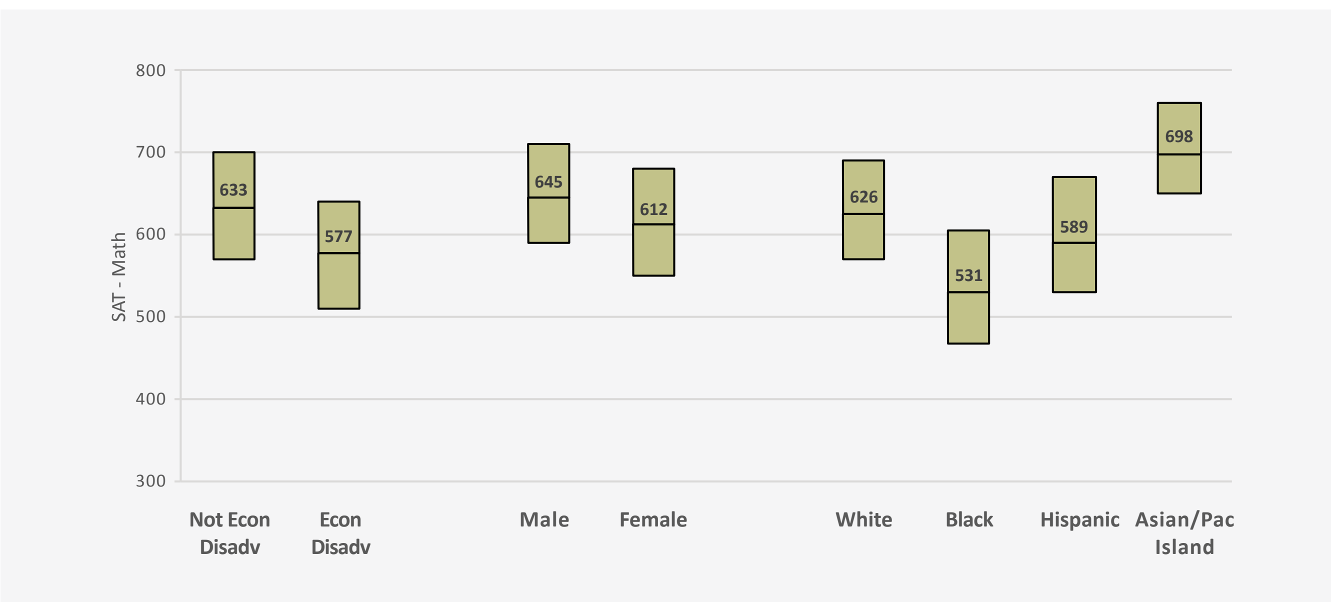
School income level is based on whether the school is above (low income) or below (high income) the median economic disadvantage rate statewide. School minority share is based on whether the school is above (high minority) or below (low minority) the median share of non-White students statewide.

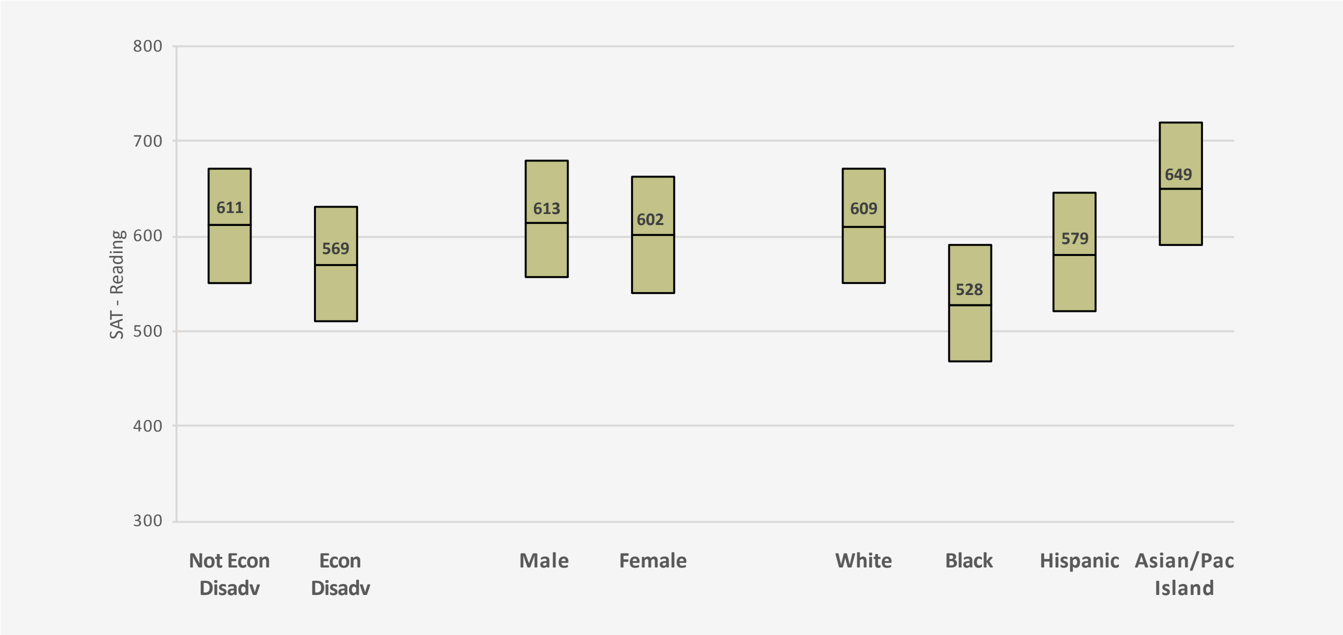
Figure A1: Estimates of demographic differences in high school achievement for third-grade high achievers



Note: Data are restricted to students in 2006–08 third-grade cohorts who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

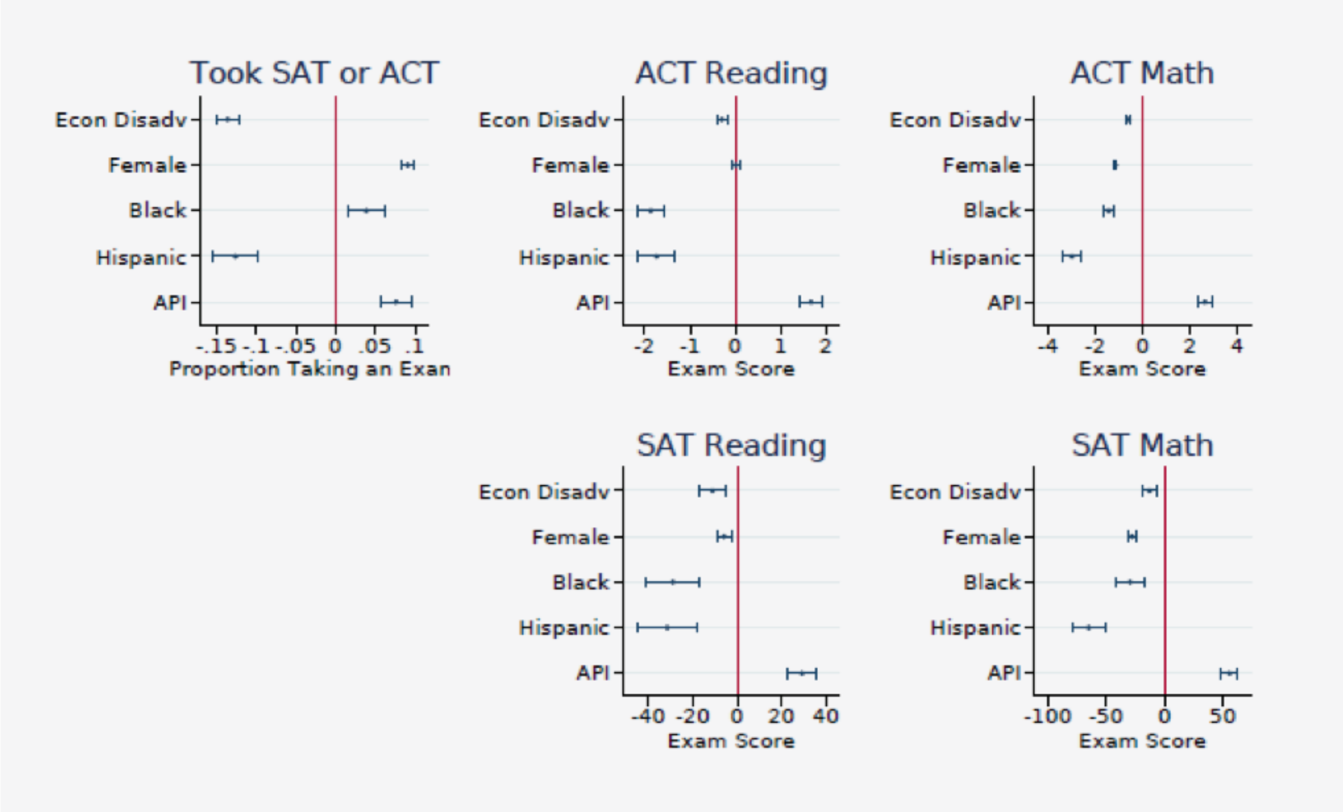
Figure A2: SAT scores for third-grade high achievers (top panel, SAT Math; bottom panel, SAT Reading)





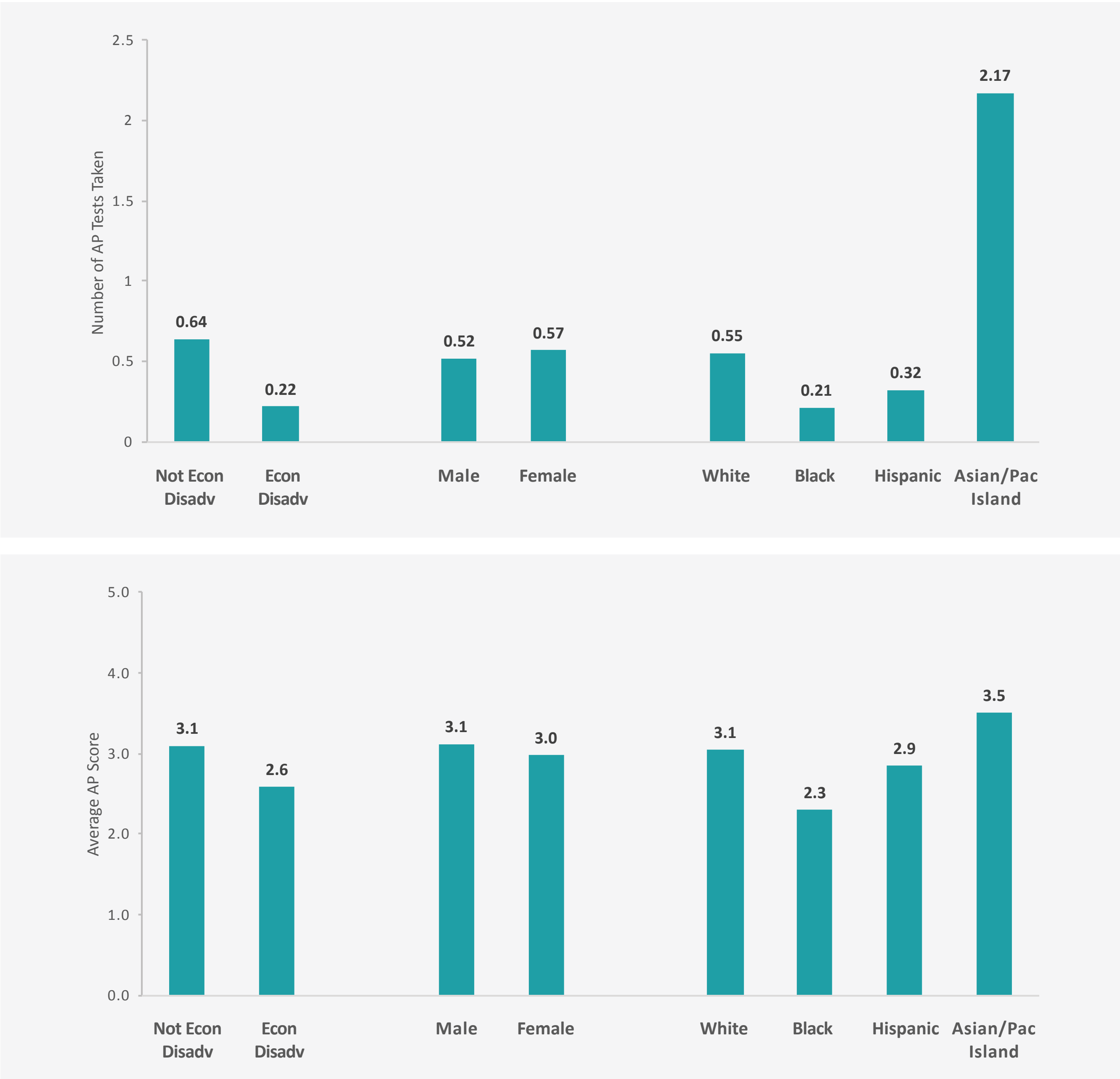
Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of the third-grade exam in same subject and are observed in an Ohio high school. Figure shows means and interquartile ranges.

Figure A3: Estimates of demographic differences in ACT/SAT exams for third-grade high achievers



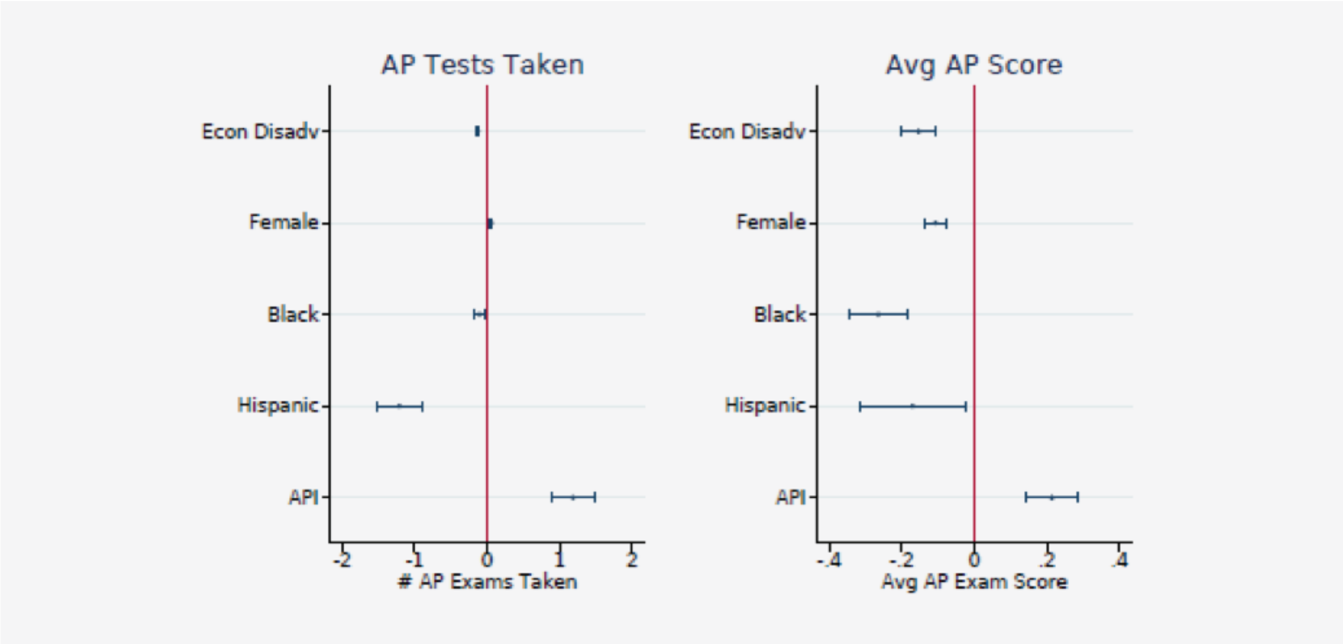
Note: Data are restricted to students in 2006–08 third-grade cohorts who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

Figure A4: AP exams for high achievers in third-grade reading (top panel, average number of AP exams taken; bottom panel, average AP-exam score)



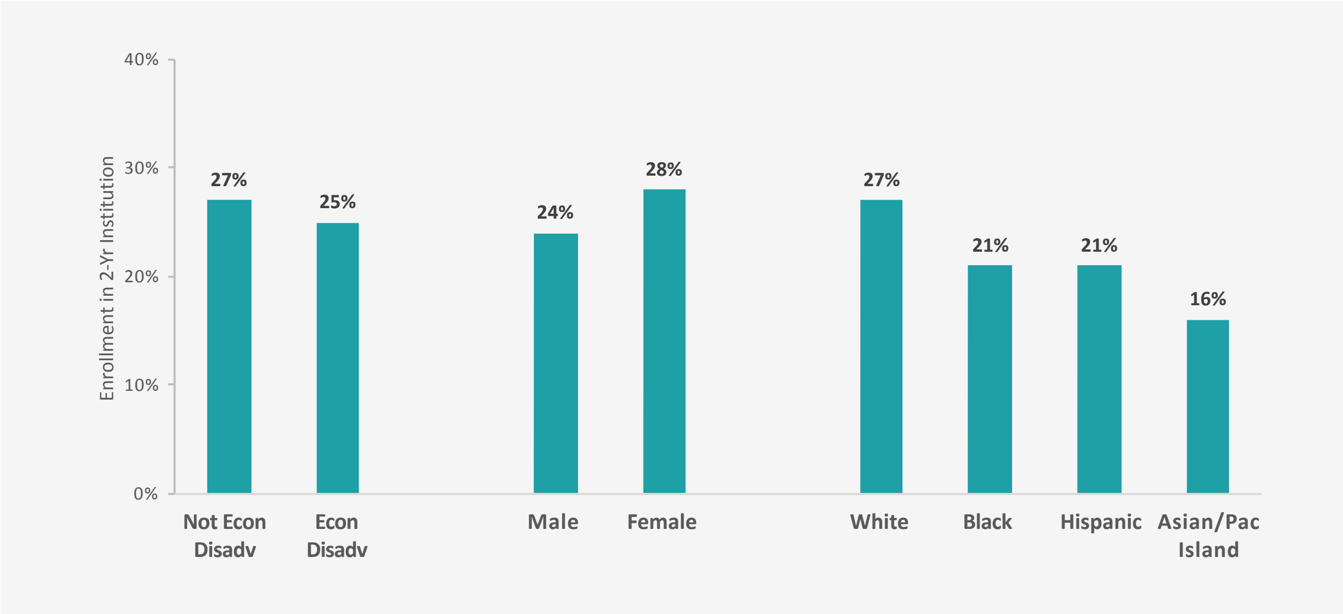
Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of the third-grade exam in math and are observed in an Ohio high school. Figure shows means.

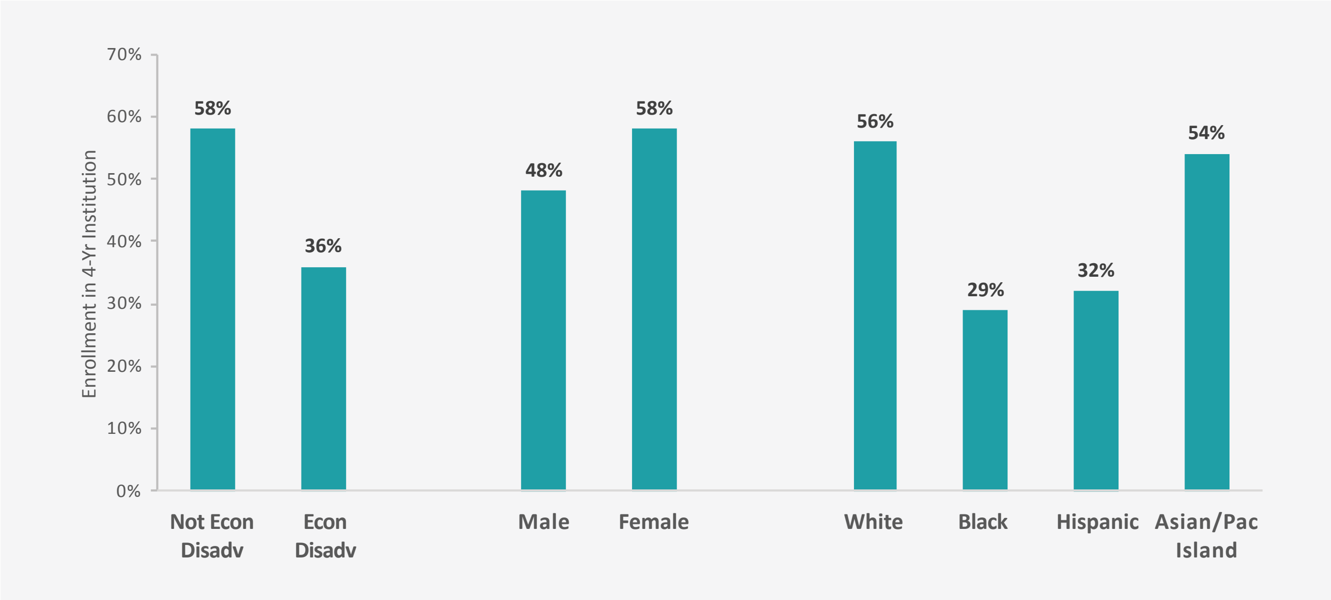
Figure A5: Estimates of demographic differences in AP testing for third-grade high achievers



Note: Data are restricted to students in 2006–08 third-grade cohorts who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

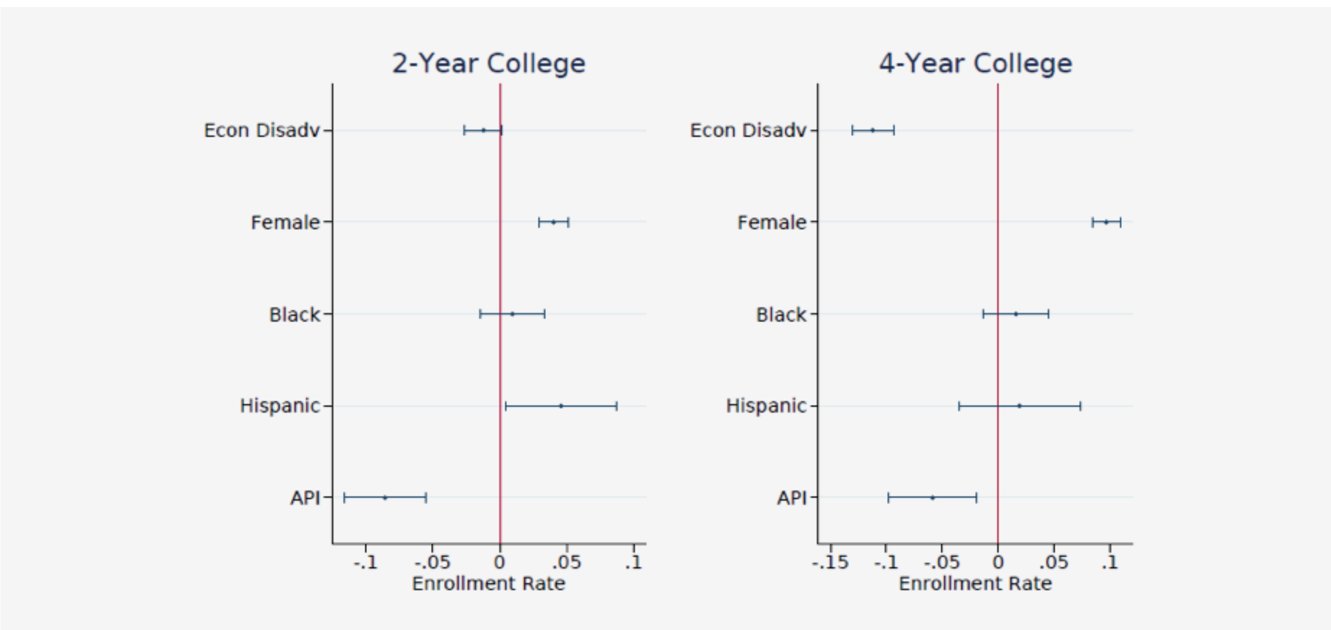
Figure A6: College enrollment for high achievers in third-grade reading (top panel, enrollment in two-year school; bottom panel, enrollment in four-year school)





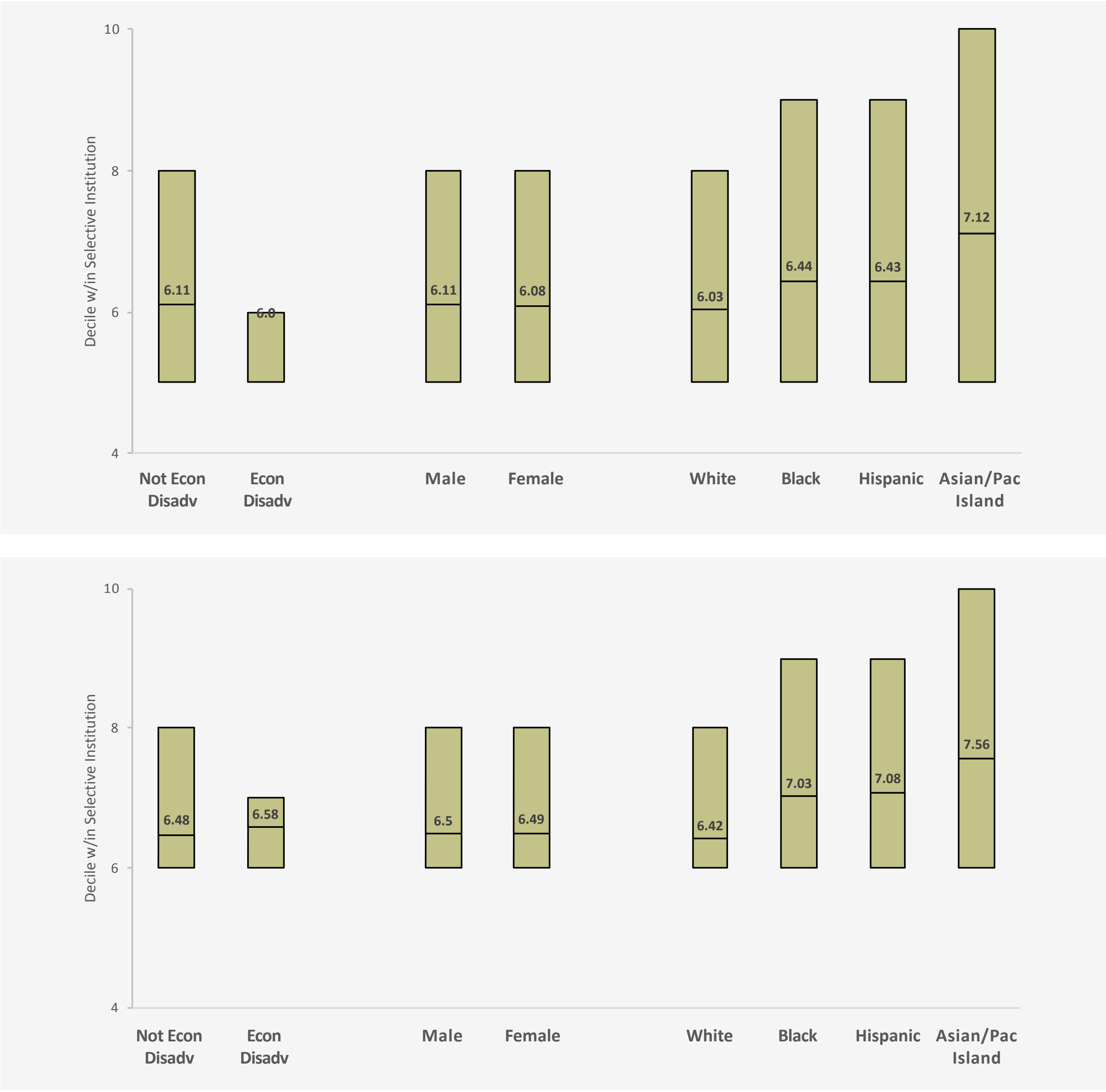
Note: Data are restricted to the 2006 third-grade cohort. Figure shows means. Enrollment based on any institution attended within two years of expected high school graduation (eleven years after third grade).

Figure A7: Estimates of demographic gaps in college enrollment for third-grade high achievers



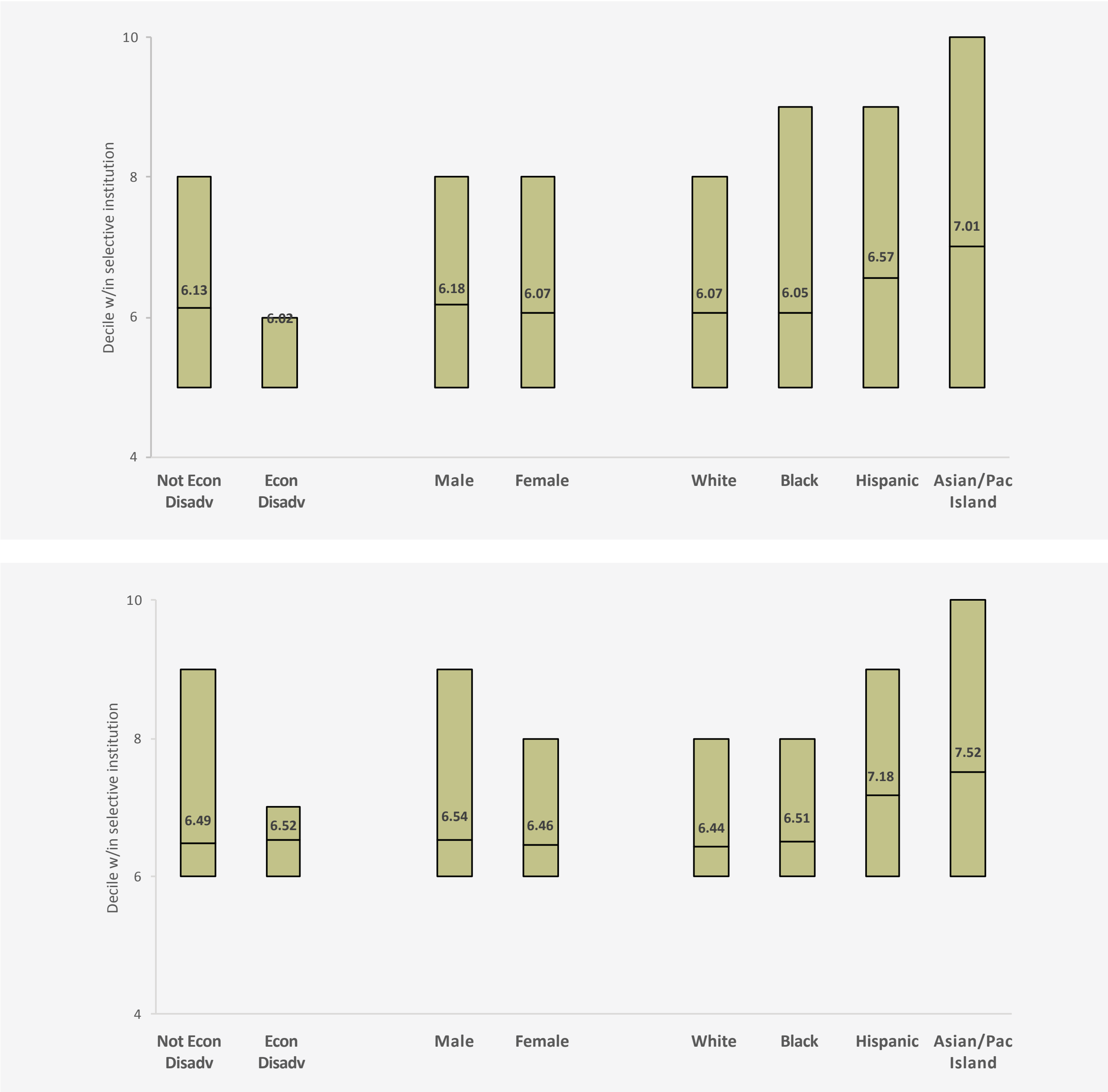
Note: Data are restricted to students in 2006 third-grade cohort who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

Figure A8: College selectivity for high achievers in third-grade math (top panel, ACT decile of college; bottom panel, SAT decile of college)



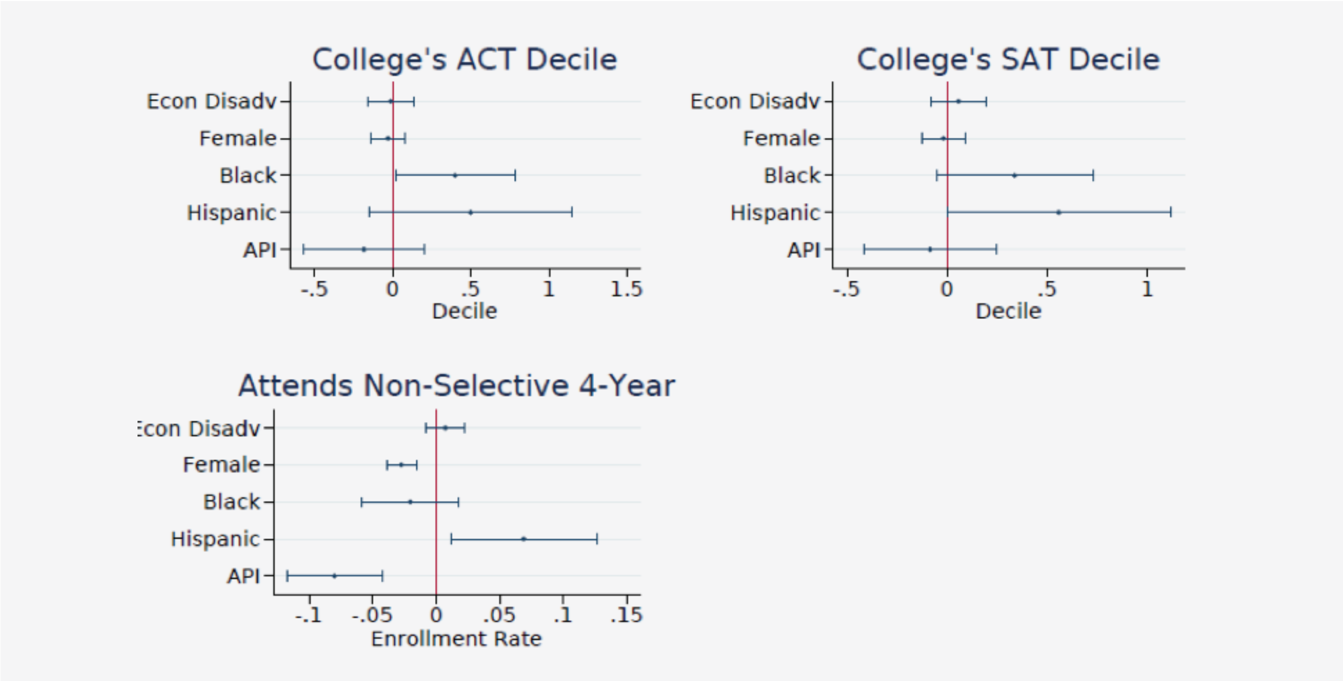
Note: Data are restricted to the 2006 third-grade cohort. Figure shows means or interquartile range. College-quality data are from IPEDS and include only those students who attended any institution within two years of expected high school graduation (eleven years after third grade).

Figure A9: College selectivity for high achievers in third-grade reading (top panel, ACT decile of college; bottom panel, SAT decile of college)



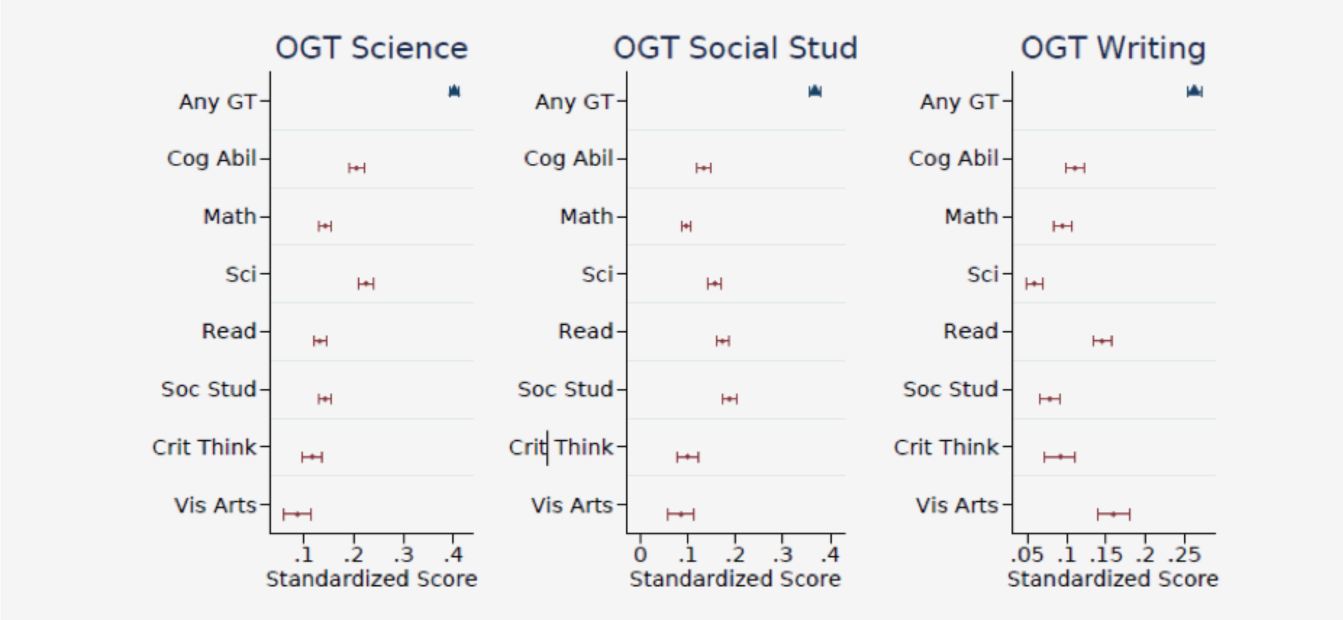
Note: Data are restricted to the 2006 third-grade cohort. Figure shows means or interquartile range. College-quality data are from IPEDS and include only those students who attended any institution within two years of expected high school graduation (eleven years after third grade).

Figure A10: Estimates of demographic gaps in college-quality measures for third-grade high achievers attending four-year colleges



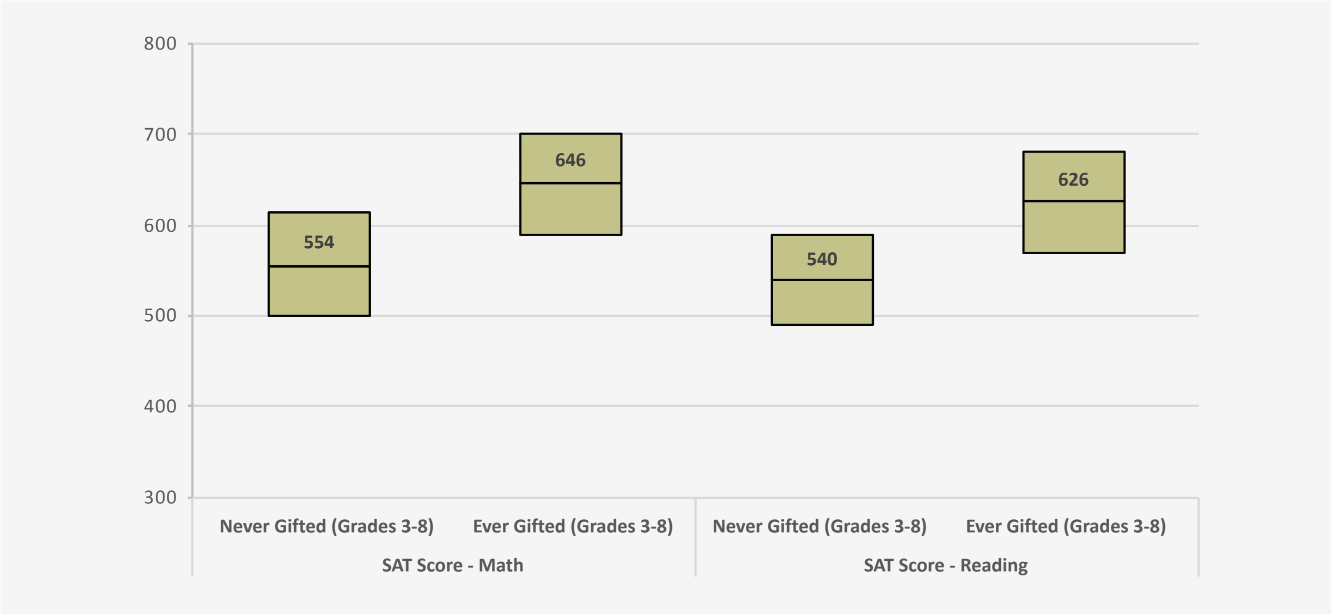
Note: Data are restricted to students in 2006 third-grade cohort who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

Figure A11: Estimates of GT differences in high school achievement with demographic controls for third-grade high achievers—OGT science, social studies, and writing



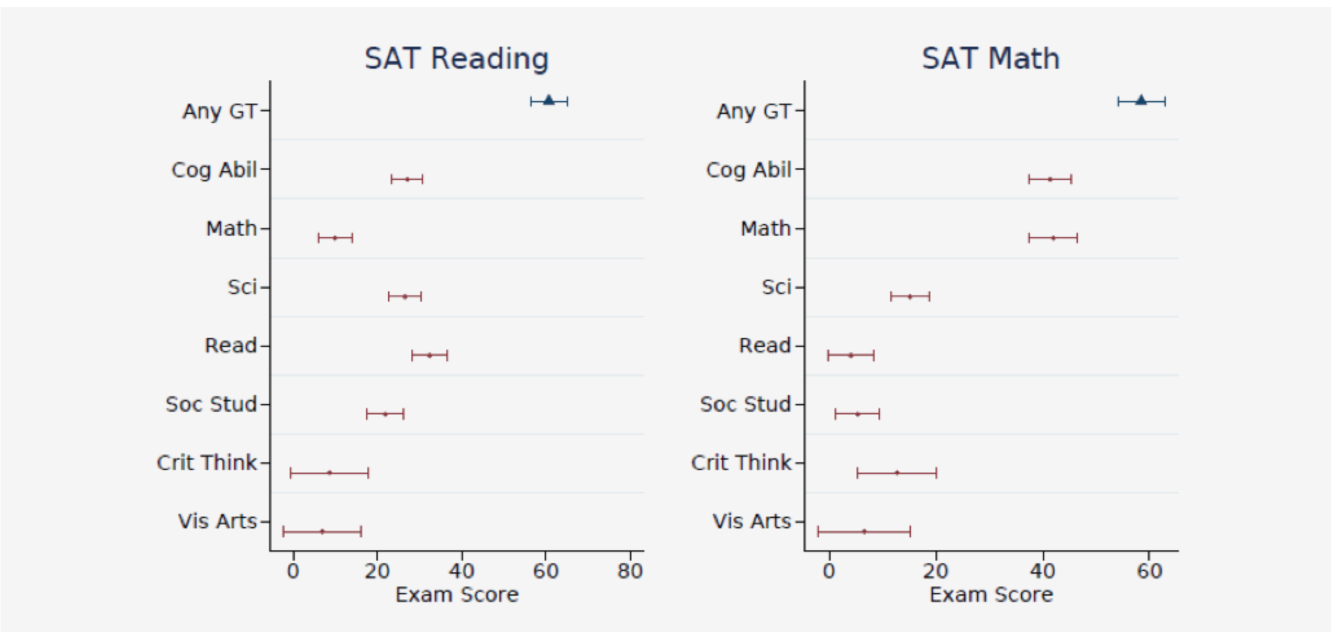
Note: Data are restricted to students in 2006–08 third-grade cohorts who score in the top quintile of math or reading in third grade. Category of GT identification listed on y-axis. Note that students can be identified in multiple categories. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects. The estimate for "Any GT" is from a separate regression.

Figure A12: SAT scores for third-grade high achievers by GT status



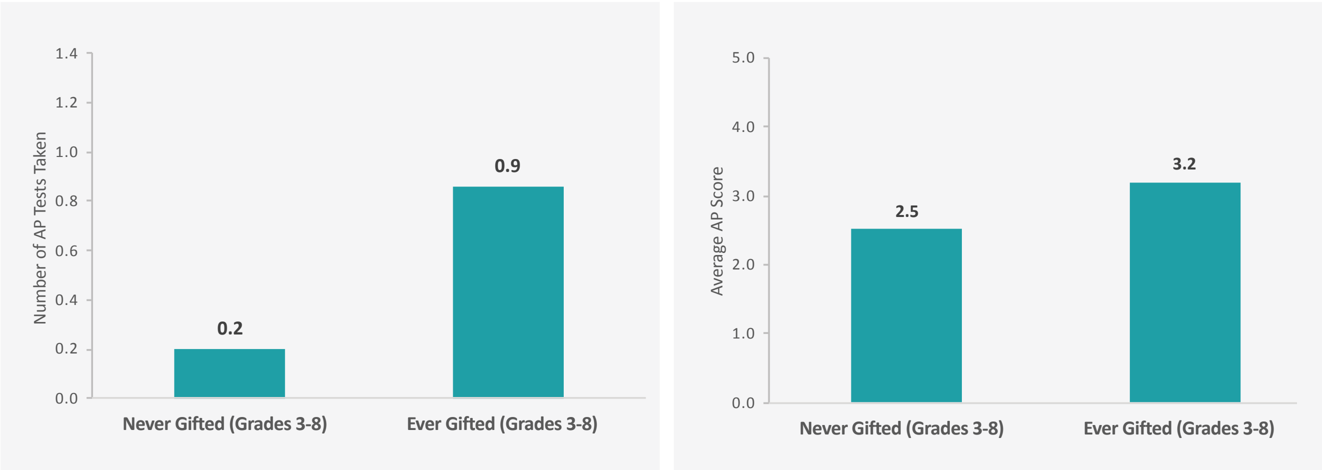
Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of the third-grade exam in same subject and are observed in an Ohio high school.

Figure A13: Estimates of GT differences in SAT exams with demographic controls for third-grade high achievers



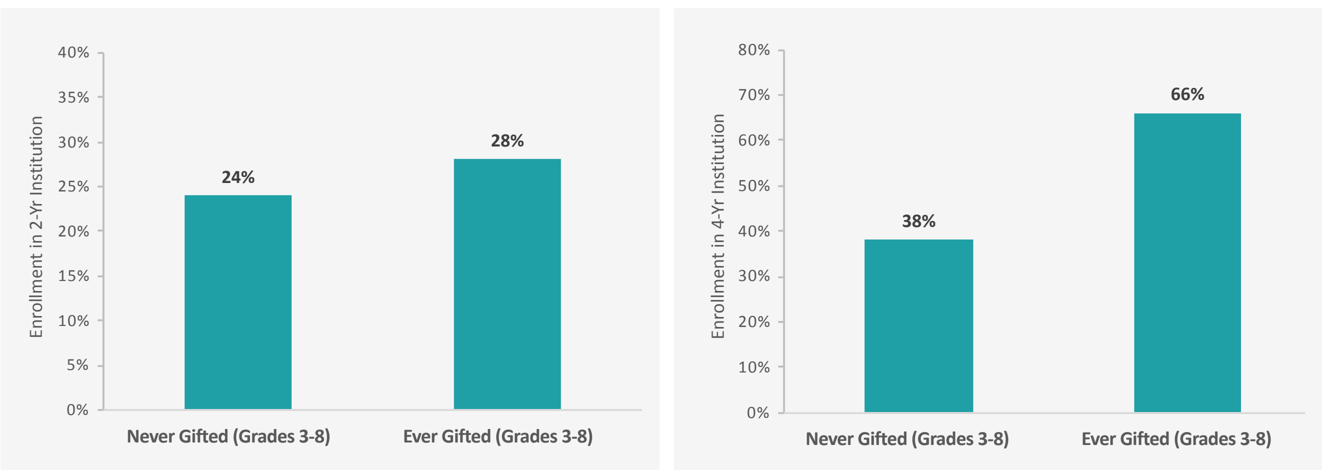
Note: Data are restricted to students in 2006–08 third-grade cohorts who score in the top quintile of math or reading in third grade. Category of GT identification listed on y-axis. Note that students can be identified in multiple categories. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects. The estimate for "Any GT" is from a separate regression.

Figure A14: AP exams for high achievers in third-grade reading by GT status



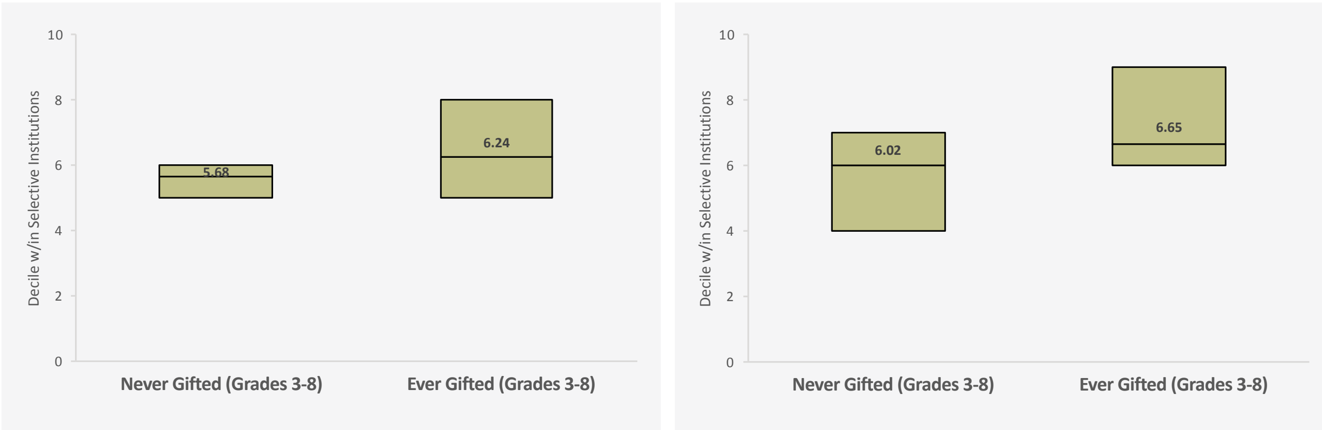
Note: Sample includes all students in 2006–08 third-grade cohorts who scored in top quintile of the third-grade exam in math and are observed in an Ohio high school. Figure shows means.

Figure A15: College enrollment for high achievers in third-grade reading by GT status



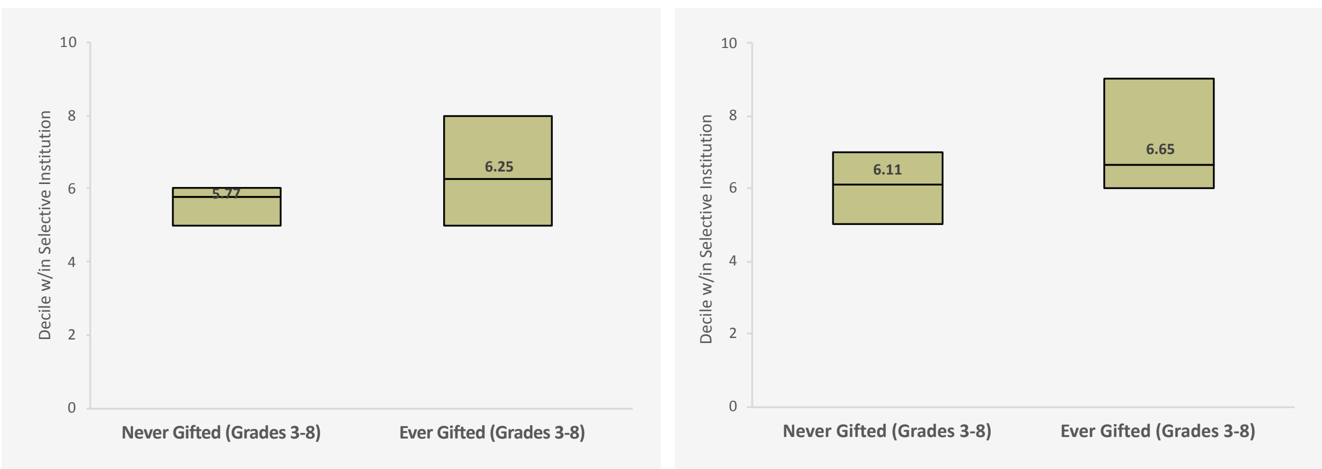
Note: Data are restricted to the 2006 third-grade cohort. Figure shows means.

Figure A16: College selectivity for high achievers in third-grade math by GT status (left panel, ACT decile of institution; right panel, SAT decile of institution)



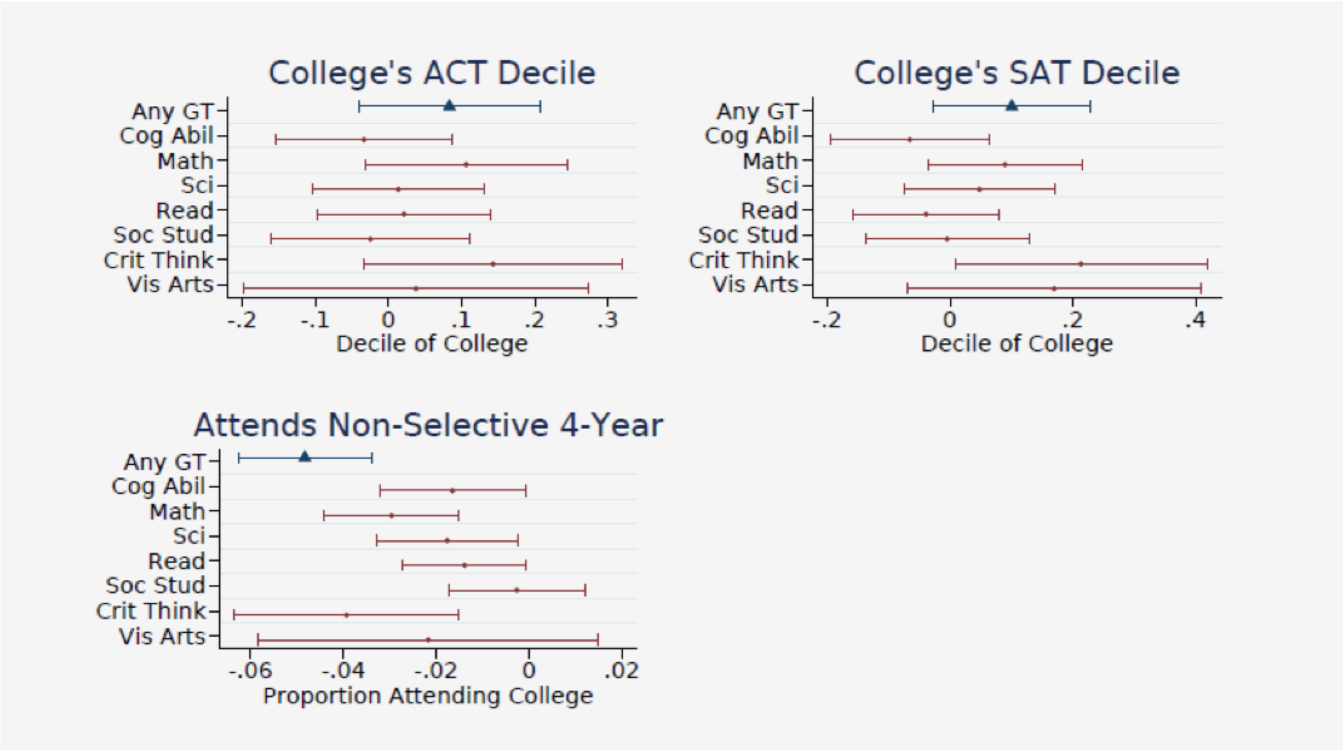
Note: Data are restricted to the 2006 third-grade cohort. Figure shows means. College-quality data from IPEDS and include only those students who attended any institution within two years of expected high school graduation (eleven years after third grade).

Figure A17: College selectivity for high achievers in third-grade reading by GT status (left panel, ACT decile of institution; right panel, SAT decile of institution)



Note: Data are restricted to the 2006 third-grade cohort. Figure shows means. College-quality data from IPEDS and include only those students who attended any institution within two years of expected high school graduation (eleven years after third grade).

Figure A18: Estimates of GT gaps in college-quality measures with demographic controls for third-grade high achievers



Note: Data are restricted to students in 2006 third-grade cohort who score in the top quintile of math or reading in third grade. Plot shows coefficient estimates and confidence intervals from a regression of the listed outcome on the demographic measures, third-grade achievement in math and reading, and school-by-cohort fixed effects.

References

Abdulkadiroğlu, Atila, Joshua Angrist, and Parag Pathak. “The elite illusion: Achievement effects at Boston and New York exam schools.” *Econometrica* 82, no. 1 (2014): 137–96.

Black, Dan A., and Jeffrey A. Smith. “Estimating the returns to college quality with multiple proxies for quality.” *Journal of Labor Economics* 24, no. 3 (2006): 701–28.

Braun, Henry, Lauren Chapman, and Sailesh Vezzu. “The Black-White achievement gap revisited.” *Education Policy Analysis Archives* 18, no. 21 (2010).

Bui, Sa A., Steven G. Craig, and Scott A. Imberman. “Is gifted education a bright idea? Assessing the impact of gifted and talented programs on students.” *American Economic Journal: Economic Policy* 6, no. 3 (2014): 30–62.

Burchinal, Margaret, Kathleen McCartney, Laurence Steinberg, Robert Crosnoe, Sarah L. Friedman, Vonnie McLoyd, Robert Pianta, and NICHD Early Child Care Research Network. “Examining the Black–White achievement gap among low‐income children using the NICHD study of early child care and youth development.” *Child Development* 82, no. 5 (2011): 1404–20.

Card, David, and Laura Giuliano. “Can tracking raise the test scores of high-ability minority students?” *American Economic Review* 106, no. 10 (2016), 2783–2816.

Card, David, and Laura Giuliano. “Universal screening increases the representation of low-income and minority students in gifted education.” *Proceedings of the National Academy of Sciences* 113, no. 48 (2016): 13678–83.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. “Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates.” *American Economic Review 104*, no. 9 (2014): 2593–632.

Chetty, Raj, and Nathaniel Hendren. “The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects.” *The Quarterly Journal of Economics* 133, no. 3 (2018): 1107–62.

Chetty, Raj, and Nathaniel Hendren. “The impacts of neighborhoods on intergenerational mobility II: County-level estimates.” *The Quarterly Journal of Economics* 133, no. 3 (2018): 1163–228.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. “The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment.” *American Economic Review* 106, no. 4 (2016): 855–902.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. “Where is the land of opportunity? The geography of intergenerational mobility in the United States.” *The Quarterly Journal of Economics* 129, no. 4 (2014): 1553–623.

Chyn, Eric. “Moved to opportunity: The long-run effects of public housing demolition on children.” *American Economic Review* 108, no. 10 (2018): 3028–56.

Cohodes, Sarah R. “The long-run impacts of specialized programming for high-achieving students.” *American Economic Journal: Economic Policy* 12, no. 1 (2020): 127–66.

Dale, Stacy B., and Alan B. Krueger. “Estimating the effects of college characteristics over the career using administrative earnings data.” *Journal of Human Resources* 49, no. 2 (2014): 323–58.

Dee, Thomas S., and Brian Jacob. “The impact of No Child Left Behind on student achievement.” *Journal of Policy Analysis and Management* 30, no. 3 (2011): 418–46.

Dobbie, Will, and Roland G. Fryer Jr. “The impact of attending a school with high-achieving peers: Evidence from the New York City exam schools.” *American Economic Journal: Applied Economics* 6, no. 3 (2014): 58–75.

Dynarski, Susan, C.J. Libassi, Katherine Michelmore, and Stephanie Owen. *Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students* (No. w25349). National Bureau of Economic Research, 2018.

Ellison, Glenn, and Ashley Swanson. “The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions.” *Journal of Economic Perspectives* 24, no. 2 (2010): 109–28.

Figlio, David N., and Helen F. Ladd. "School accountability and student achievement." In *Handbook of Research in Education Finance and Policy, 2nd Edition*, edited by Helen F. Ladd and Margaret E. Goertz, 166–82. New York: Routledge, 2015.

Ford, Donna Y. "The underrepresentation of minority students in gifted education: Problems and promises in recruitment and retention." *The Journal of Special Education* 32, no. 1 (1998): 4–14.

Fryer Jr, Roland G., and Steven D. Levitt. "Understanding the Black-White test score gap in the first two years of school." *Review of Economics and Statistics* 86, no. 2 (2004): 447–64.

Gaddis, S. Michael, and Douglas Lee Lauen. "School accountability and the black-white test score gap." *Social Science Research* 44 (2014): 15–31.

Grissom, Jason A., and Christopher Redding. "Discretion and disproportionality: Explaining the underrepresentation of high-achieving students of color in gifted programs." *Aera Open* 2, no. 1 (2016).

Hanushek, Eric A., and Margaret E. Raymond. "The effect of school accountability systems on the level and distribution of student achievement." *Journal of the European Economic Association* 2, no. 2–3 (2004): 406–15.

Hanushek, Eric A., and Steven G. Rivkin. "Harming the best: How schools affect the Black-White achievement gap." *Journal of Policy Analysis and Management* 29, no. 3 (2009): 366–93.

Harris, Douglas N., and Carolyn D. Herrington. "Accountability, standards, and the growing achievement gap: Lessons from the past half-century." *American Journal of Education* 112, no. 2 (2006): 209–38.

Hoxby, Caroline M., and Christopher Avery. "Missing 'One-Offs': The Hidden Supply of High-Achieving, Low-Income Students." *Brookings Papers on Economic Activity* 1 (2013): 1–46.

Hoxby, Caroline M., and Sarah Turner. "What high-achieving low-income students know about college." *American Economic Review: Papers and Proceedings* 105, no. 5 (2015): 514–17.

Jackson, C. Kirabo. "A little now for a lot later: A look at a Texas advanced placement incentive program." *Journal of Human Resources* 45, no. 3 (2010): 591–639.

Jacob, Brian A. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89, no. 5–6 (2005): 761–96.

Jencks, Christopher, and Meredith Phillips, eds. *The Black-White test score gap*. Washington, D.C.: Brookings Institution Press, 1998.

Joseph, Laurice M., and Donna Y. Ford. "Nondiscriminatory assessment: Considerations for gifted education." *Gifted Child Quarterly* 50, no. 1 (2006): 42–51.

Lee, Jaekyung Lee, and Kenneth K. Wong. "The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes." *American Educational Research Journal* 41, no. 4 (2004): 797–832.

Long, Mark C. "College quality and early adult outcomes." *Economics of Education Review* 27, no. 5 (2008): 588–602.

Neal, Derek, and Diane Whitmore Schanzenbach. "Left behind by design: Proficiency counts and test-based accountability." *The Review of Economics and Statistics* 92, no. 2 (2010): 263–83.

Reardon, Sean F., and Claudia Galindo. "The Hispanic-White achievement gap in math and reading in the elementary grades." *American Educational Research Journal* 46, no. 3 (2009): 853–91.

Reardon, Sean F., Joseph P. Robinson-Cimpian, and Ericka S. Weathers. "Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps." In *Handbook of Research in Education Finance and Policy, 2nd Edition*, edited by Helen F. Ladd and Margaret E. Goertz, 497–516. New York: Routledge, 2015.

Redding, Christopher, and Jason A. Grissom. "Do students in gifted programs perform better? Linking gifted program participation to achievement and nonachievement outcomes." *Educational Evaluation and Policy Analysis* (forthcoming). doi:10.3102/01623737211008919.

Xiang, Yun, Michael Dahlin, John Cronin, Robert Theaker, and Sarah Durant. *Do high flyers maintain their altitude? Performance trends of top students*. Washington, D.C.: Thomas B. Fordham Institute, 2011.